Yield gene cloning in wheat

Liuling Yan

Department of Plant and Soil Sciences
Oklahoma State University
Stillwater, OK, U.S.A.

Outline

1. **The wheat CAP projects**
    1.1. The wheat CAP3 project
    1.2. A major QTL *QYld.osu-1BS* for grain yield
    1.3. *TaCol-B5* for spike architecture and grain yield
2. **Mapping and positional cloning of yield genes**
    2.1. Parental lines used for mapping and cloning
    2.2. Natural *Q5A* variants in the cloning population
    2.3. Mapping of *QSns.osu-7B*
    2.4. Cloning and validation of *TaCol-B5*
3. **Post-cloning activities on *TaCol-B5***
    3.1. Field validation
    3.2. Mechanisms of the cloned gene underlying the traits
    3.3. Origin and distribution of *TaCol-B5*
4. **Perspectives**

1. The wheat CAP projects

1.1. The wheat CAP3 project

https://www.triticeaecap.org/qtl-cloning-projects

WHAT'S INSIDE

- Home
- About WheatCAP
  - Participants Directory
  - Scientific Advisory Board
  - Industry Liaison Committee
  - Education Advisory Board
  - Terminology
- QTL cloning projects
- The WheatCAP Education Team
- TCAP Production (2011-2016)

FORUM – RECENT POSTS

- Forums

USDA United States Department of Agriculture National Institute of Food and Agriculture

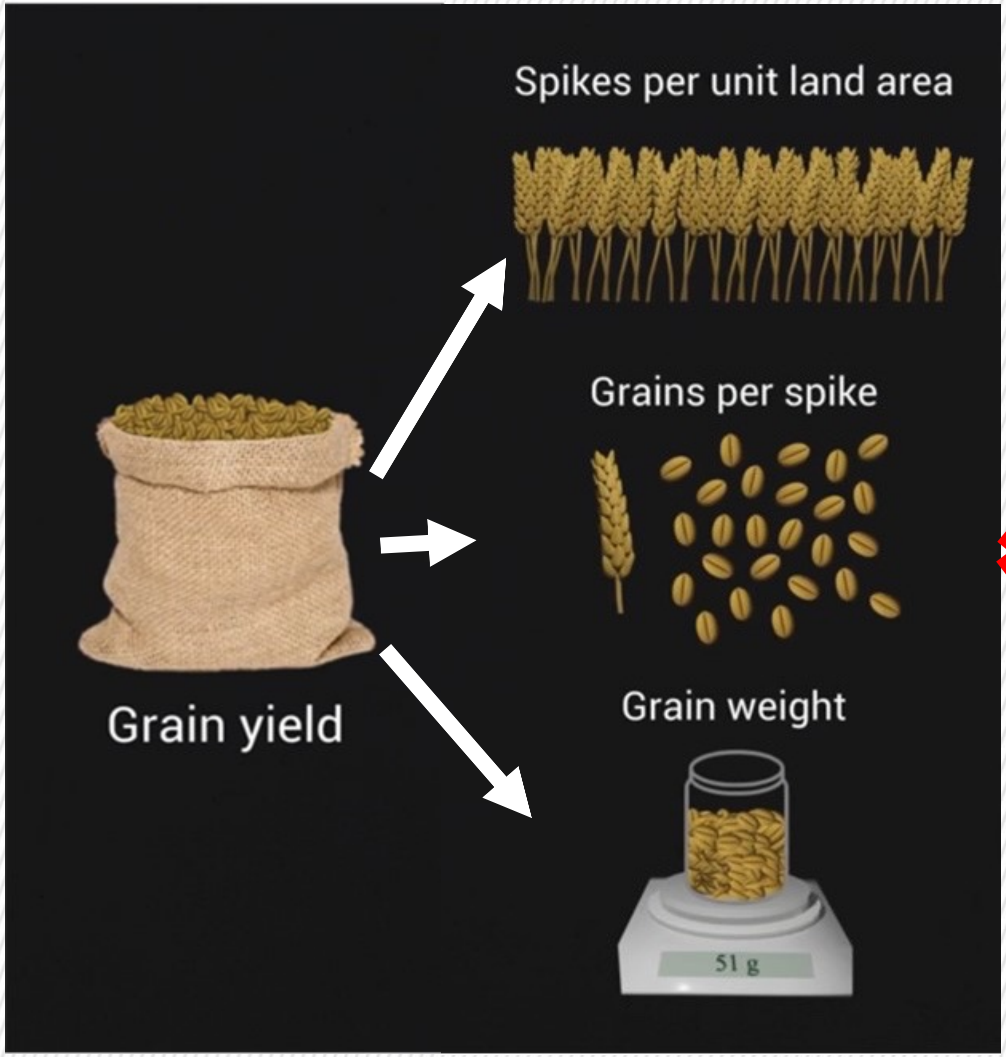
QTL cloning projects

You can find the QTL cloning projects every group proposed along with contact information. These pages will be updated as the groups move forward.

- University of California, Davis (spikelets per spike)
- Colorado State University (grain width, thousand kernel weight)
- Cornell University (seed size)
- University of Idaho (spikelets per spike)
- QTL cloning (tiller number) and training activities at Montana State University
- Texas A&M AgriLife Research (yield, thousand kernel weight and harvest index)
- USDA-ARS, Fargo, ND (grain yield)
- USDA-ARS and Kansas State University Collaborative Project (yield)
- South Dakota State University (yield)
- University of Minnesota (kernel weight and kernel width)
- Oklahoma State University (yield)
- North Carolina State University and USDA-ARS, Raleigh, NC (kernel weight)
- University of Arkansas (yield, thousand-kernel weight)
- Washington State University (spike length, kernels per spike and thousand kernel weight)

Dr. Jorge Dubcovsky
UC Davis

Dissection of wheat grain yield

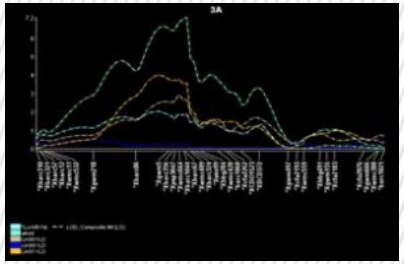


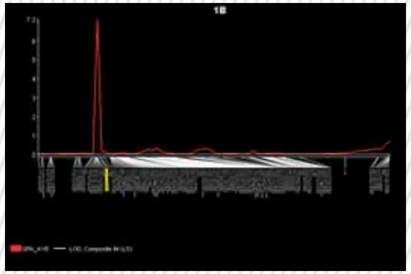
Three components of grain yield

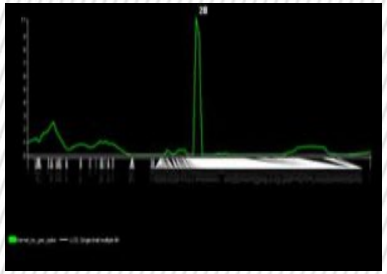
Two subcomponents of grains per spike

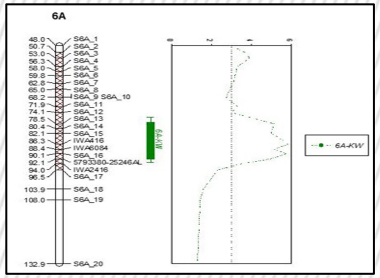
Yield QTLs in the wheat CAP3

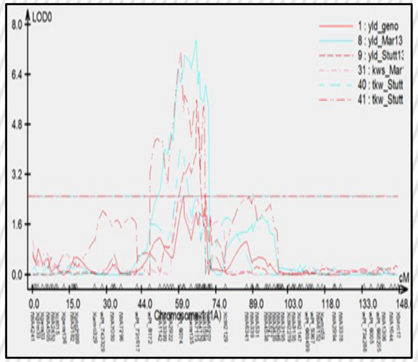
1	University of California, Davis	Spikelets per spike	7AL (160-165 Mb)
2	Colorado State University	Grain width, thousand kernel weight	6BL
3	Cornell University	Seed size	5BS
4	University of Idaho	Spikelets per spike	5AL
5	Montana State University	Tiller number	6BS (30 Mb)
6	Texas A&M AgriLife Research	Yield, thousand kernel weight and harvest index	2BS (18 Mb)
7	USDA-ARS, Fargo, ND	Grain yield, Spikelets per spike, grains per spike	3A, 1BS, 2BS
8	USDA-ARS and Kansas State University	Yield	2DL (111 Mb)
9	South Dakota State University	Yield	7DS
10	University of Minnesota	Kernel weight and kernel width	2A
11	**Oklahoma State University**	**Yield**	**1BS**
12	North Carolina State University and USDA-ARS	Kernel weight	6A
13	University of Arkansas	Yield, thousand-kernel weight	1A, 3AL
14	Washington State University	Spike length, kernels per spike and thousand kernel weight	4AL
15	Michigan State University	Yield	2DL (637-640 Mb)

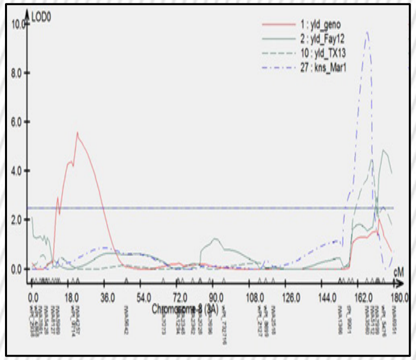


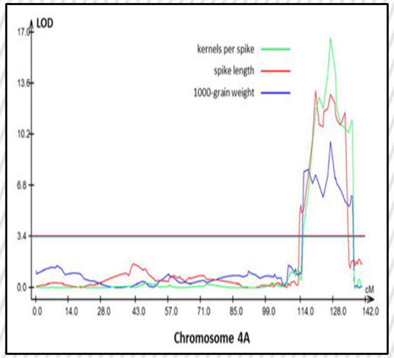


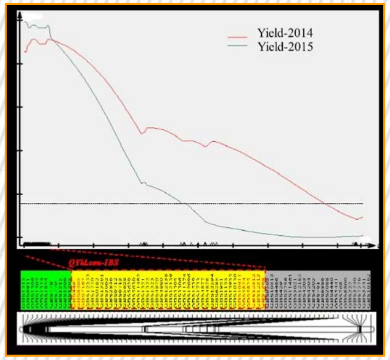












1.2. A major QTL *QYld.osu-1BS* for grain yield

(Initially proposed gene cloning project in Oklahoma)

Results:

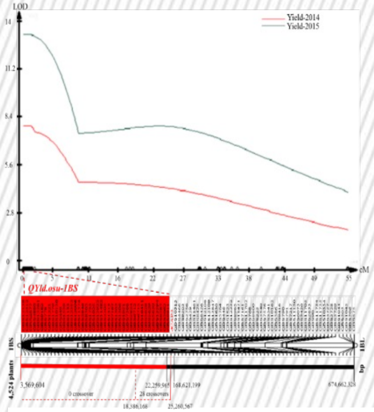
- Duster and Billings, two locally adapted winter wheat cultivars.
- 260 doubled haploid (DH) lines.
- GBS markers for the DH population.
- Field-based grain yield from breeding plots for two years.
- A major QTL on chromosome arm 1BS (*QYld.osu-1BS*) for 13.9% and 23.5% of the total phenotypic variation.
- No crossover in the targeted region from the telomere to 18.4 Mb among >4,000 individual $F_{2:4}$ plants.
- KASP markers for unique sequences in the *QYld.osu-1BS* region in the Duster allele.

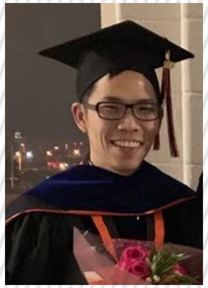
Conclusion:

- *QYld.osu-1BS* could not be cloned in a short term.
- Three KASP markers can be used to introduce the higher yield allele in Duster into new wheat variety.

TaHf1, Li *et al.*, BMC Genomics. 16, (2015)

TaOGT1, Fan *et al.*, Nat Commun. 12, (2021)





Forrest Kan

Mol Breeding (2020) 40: 26
https://doi.org/10.1007/s11032-020-1106-0

Genetic characterization and deployment of a major gene for grain yield on chromosome arm 1BS in winter wheat

Chia-Cheng Kan · Haiyan Jia · Carol Powers · Brett F. Carver · Liuling Yan

Kan *et al*., Mol. Breed. 40: 26 (2020).

1.3. *TaCol-B5* for spike architecture and grain yield

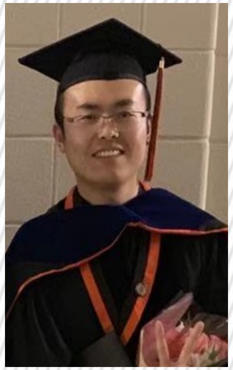
Science AAAS

RESEARCH

PLANT SCIENCE

TaCol-B5 modifies spike architecture and enhances grain yield in wheat

Xiaoyu Zhang[1]†, Haiyan Jia[1,2]†, Tian Li[1,3]†, Jizhong Wu[1,4], Ragupathi Nagarajan[1], Lei Lei[1],
Carol Powers[1], Chia-Cheng Kan[1], Wei Hua[5], Zhiyong Liu[6], Charles Chen[7],
Brett F. Carver[1]*, Liuling Yan[1]*



Xiaoyu Zhang

Dr. Haiyan Jia

Dr. Tian Li

Science 376, 180-183 (2022)

https://www.science.org/doi/10.1126/science.abm0717

2. Mapping and positional cloning of yield genes

2.1. Parental lines used for gene cloning

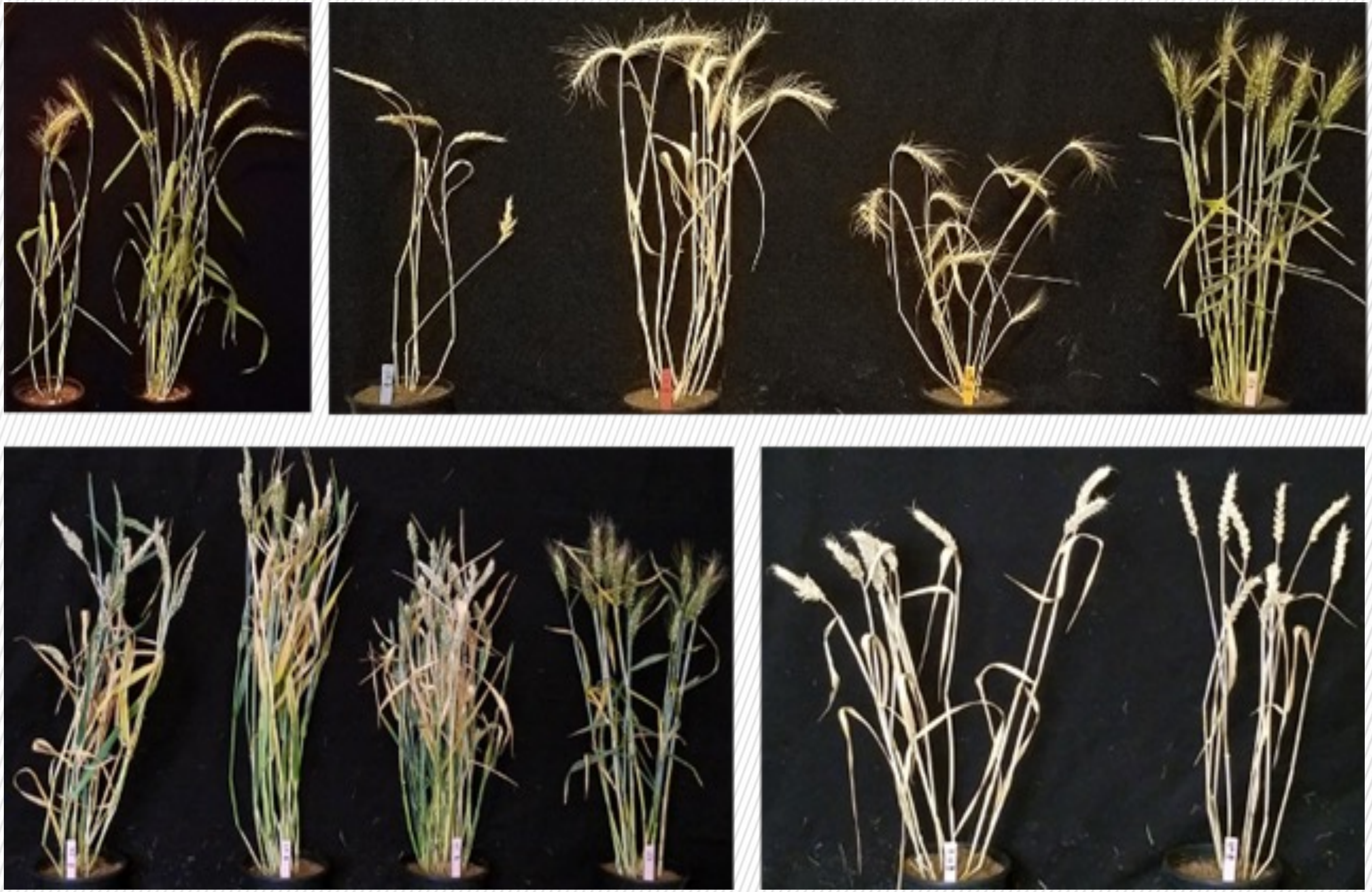
CItr 17600 (L20) Yangmai18 (Y18)

CItr 17600 (L20) Yangmai18 (Y18)

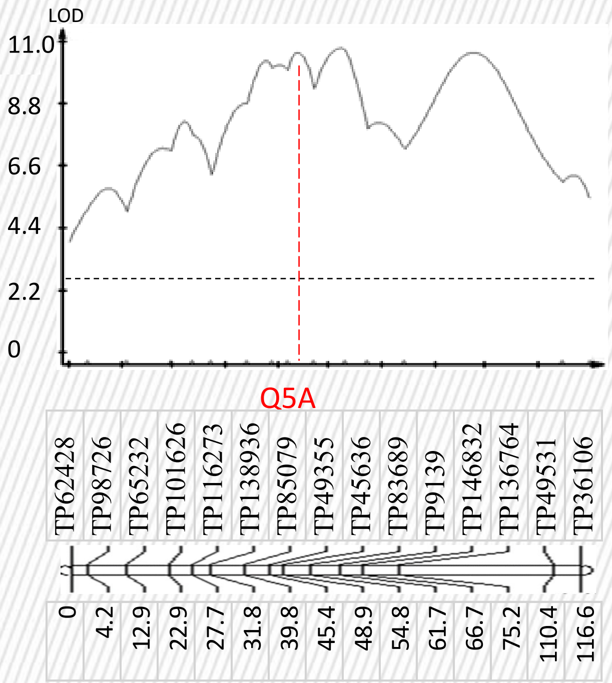
CIMMYT

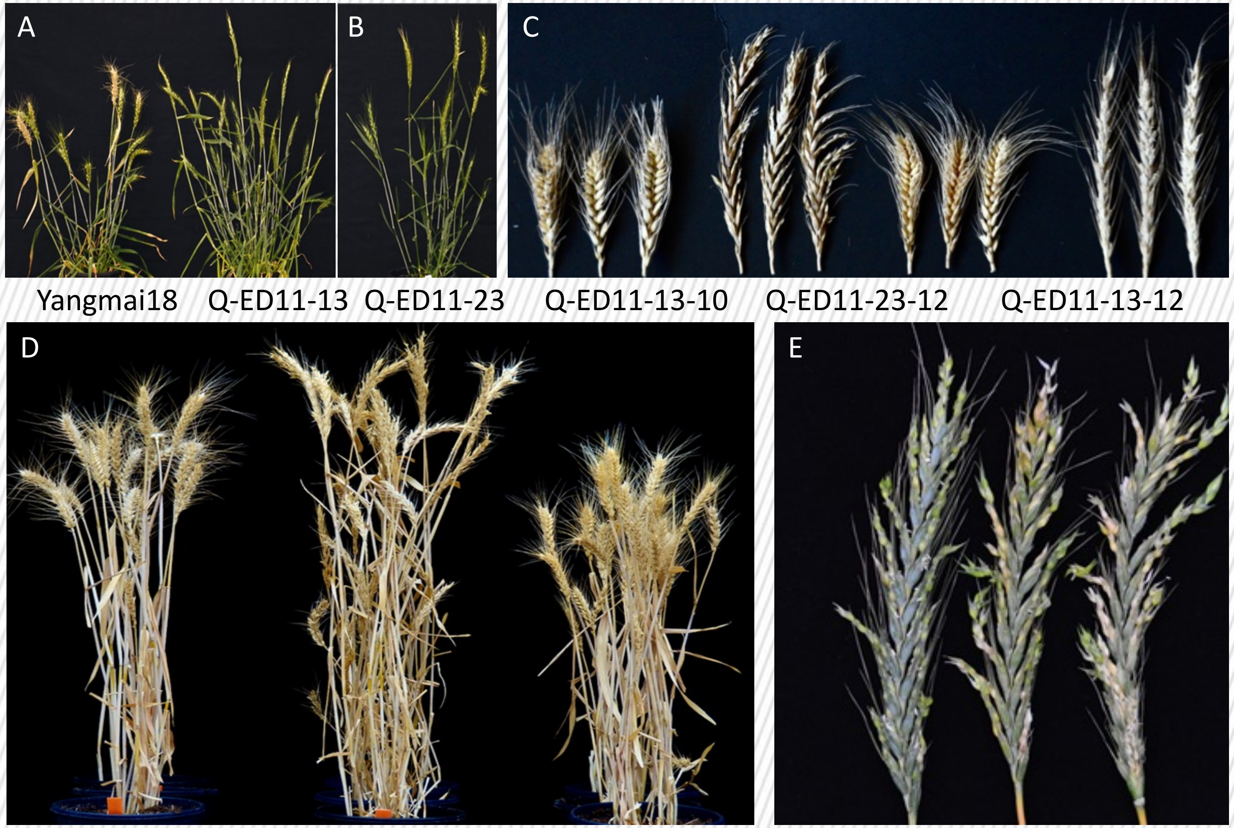
Yangzhou

Diversity and segregation of plant/spike traits in the L20 x Y18 F$_{2:3}$ populations



2.2. Natural *Q5A* variants in the cloning population





Yangmai18 Q-ED11-13 Q-ED11-23 Q-ED11-13-10 Q-ED11-23-12 Q-ED11-13-12

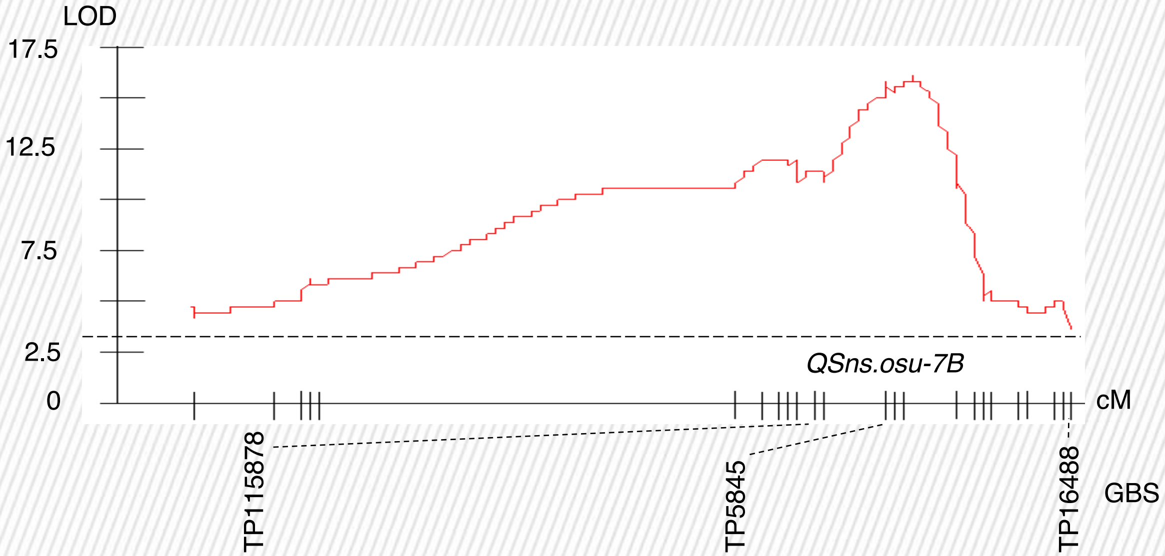
Yangmai18 Q-ED11-23-12 Q-ED11-13-18 Q-ED11-13-17

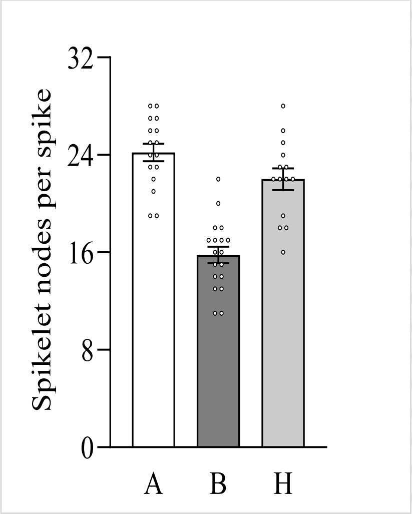
Min Fan

Natural variants and editing events provide insights into routes for spike architecture modification in common wheat

Fan *et al*., The Crop Journal, 2022.

2.3. Mapping of *QSns.osu-7B*





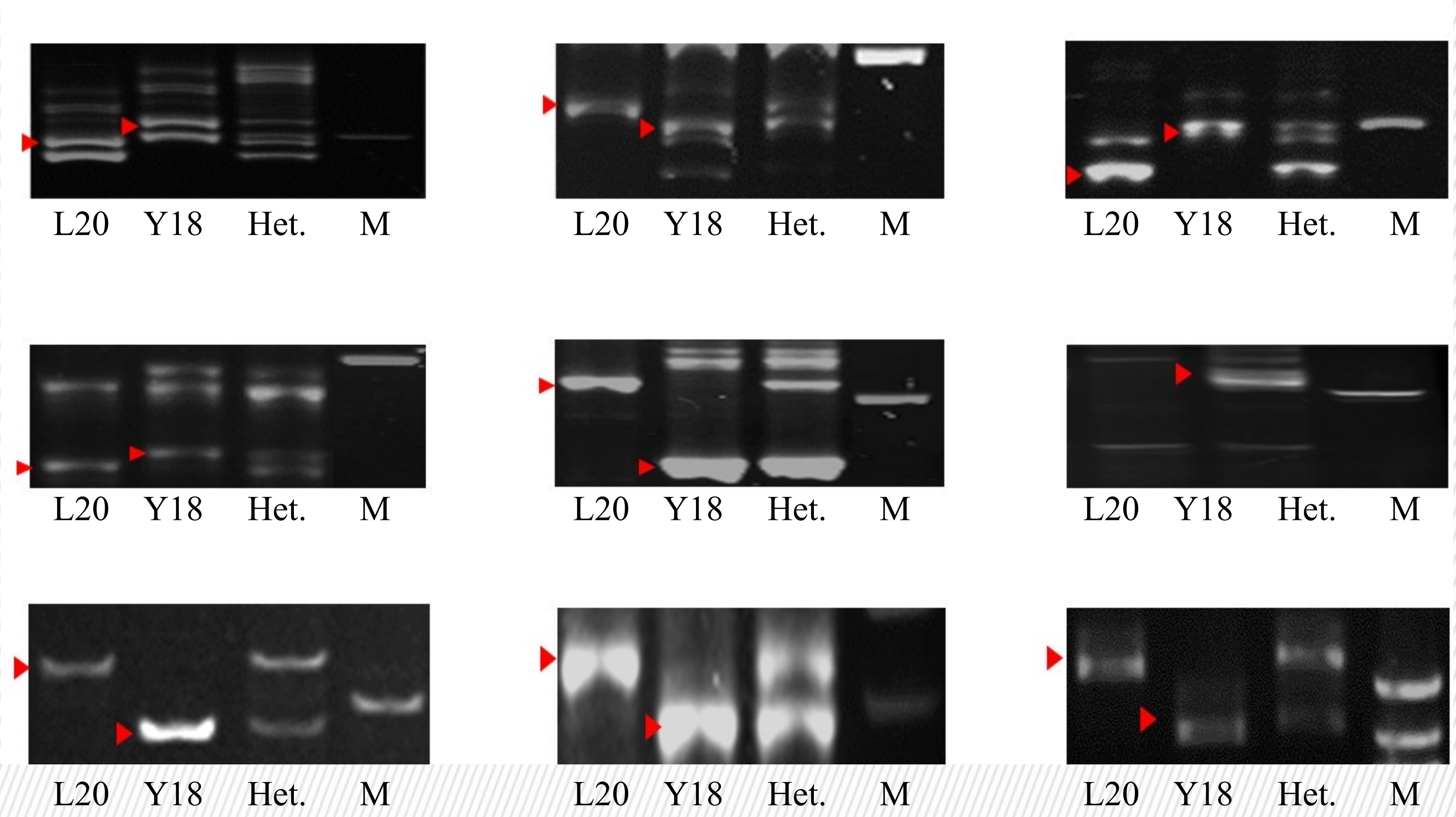
A major QTL for the number of spikelet nodes per spike

QSns.osu-7B on chromosome arm 7BL

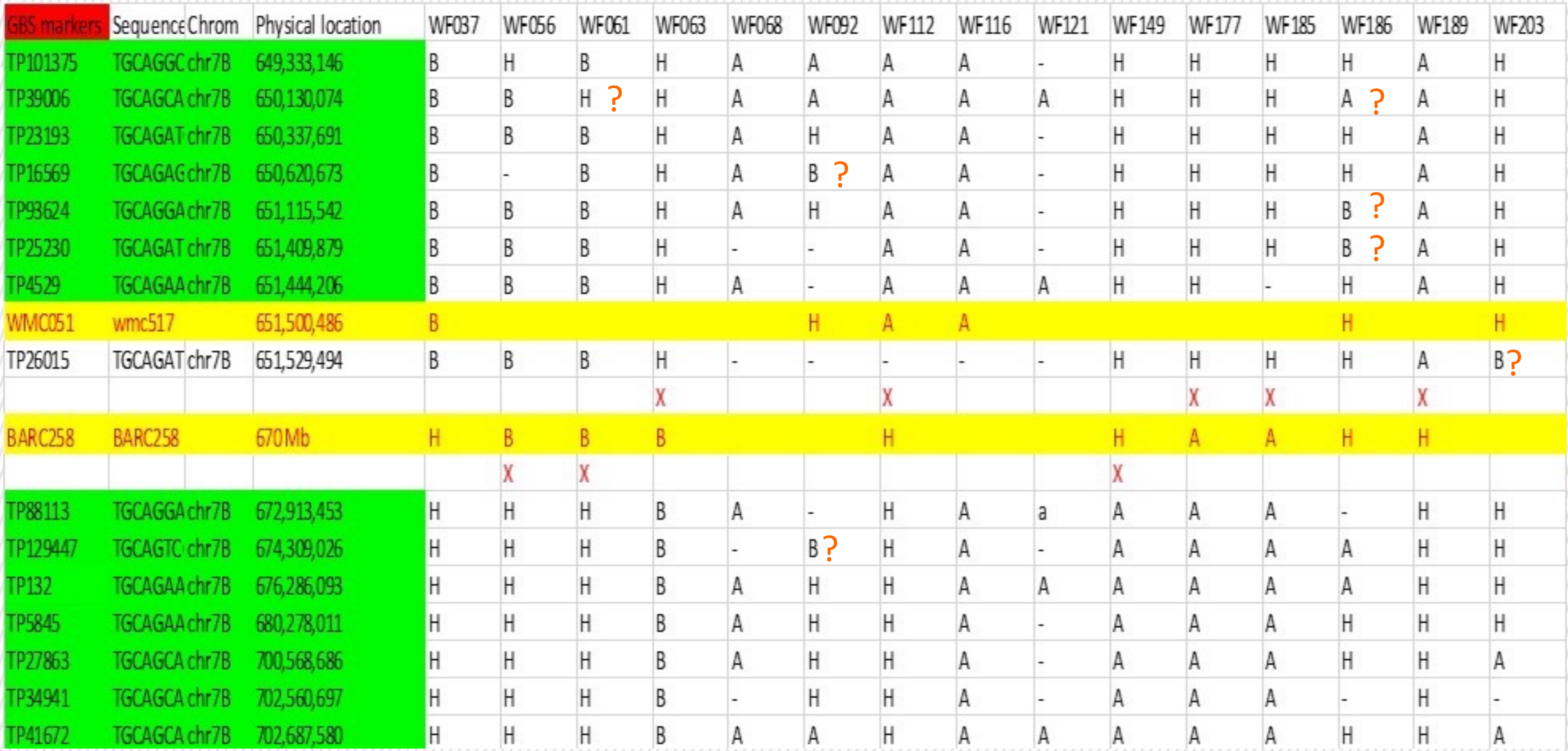
LOD value:15.3

Effect: 43.1% of the total phenotypic variation.

Molecular markers for fine map of *QSns.osu-7B*



Recombinant lines at the *QSns.osu-7B* locus



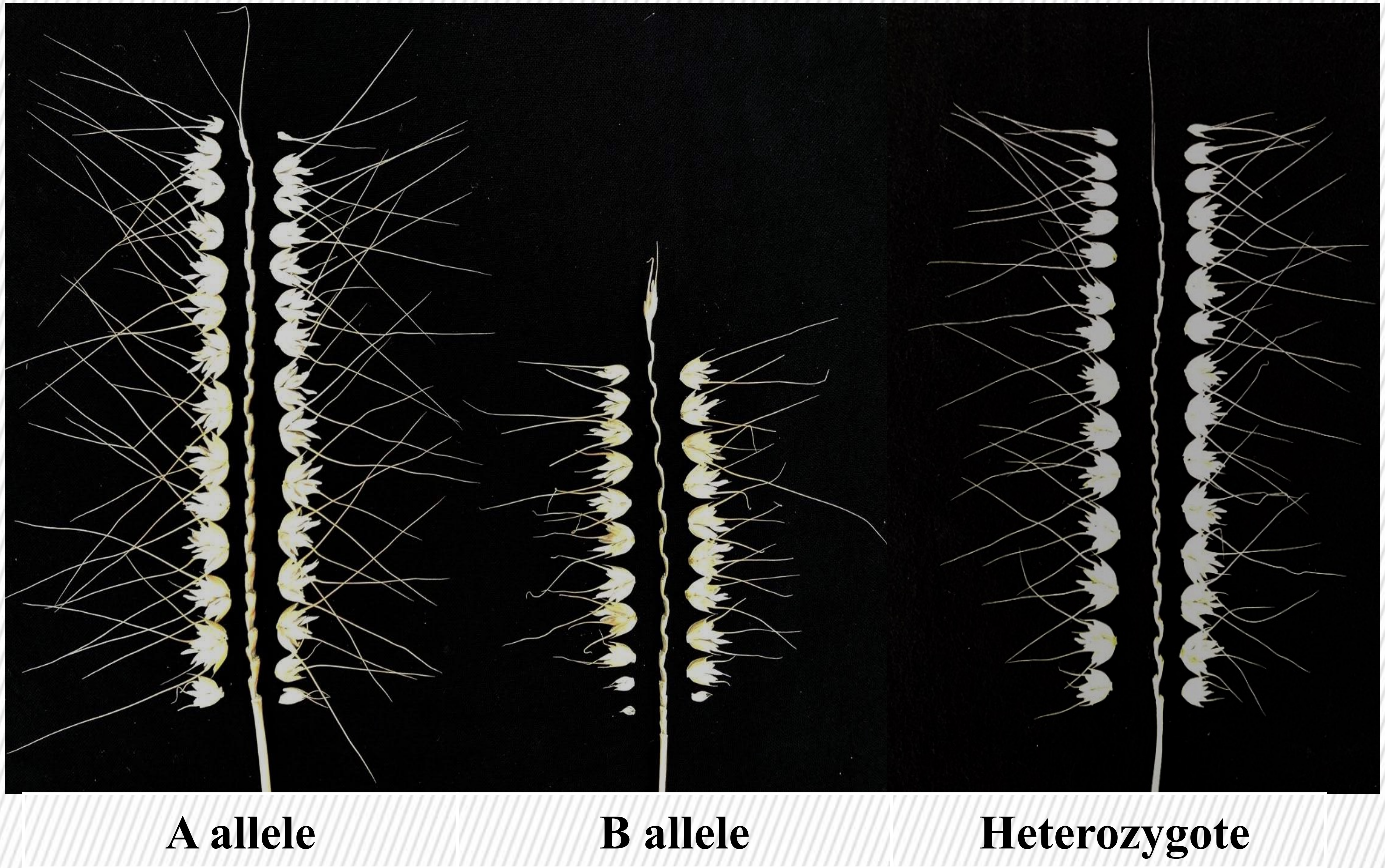
GBS markers	Sequence	Chrom	Physical location	WF037	WF056	WF061	WF063	WF068	WF092	WF112	WF116	WF121	WF149	WF177	WF185	WF186	WF189	WF203
TP101375	TGCAGGC	chr7B	649,333,146	B	H	B	H	A	A	A	A	-	H	H	H	H	A	H
TP39006	TGCAGCA	chr7B	650,130,074	B	B	H ?	H	A	A	A	A	A	H	H	H	A ?	A	H
TP23193	TGCAGAT	chr7B	650,337,691	B	B	B	H	A	H	A	A	-	H	H	H	H	A	H
TP16569	TGCAGAG	chr7B	650,620,673	B	-	B	H	A	B ?	A	A	-	H	H	H	H	A	H
TP93624	TGCAGGA	chr7B	651,115,542	B	B	B	H	A	H	A	A	-	H	H	H	B ?	A	H
TP25230	TGCAGAT	chr7B	651,409,879	B	B	B	H	-	-	A	A	-	H	H	H	B ?	A	H
TP4529	TGCAGAA	chr7B	651,444,206	B	B	B	H	A	-	A	A	A	H	H	-	H	A	H
WMC051	wmc517		651,500,486	B					H	A	A					H		H
TP26015	TGCAGAT	chr7B	651,529,494	B	B	B	H	-	-	-	-	-	H	H	H	H	A	B?
							X			X				X	X		X	
BARC258	BARC258		670 Mb	H	B	B	B			H			H	A	A	H	H	
					X	X							X					
TP88113	TGCAGGA	chr7B	672,913,453	H	H	H	B	A	-	H	A	a	A	A	A	-	H	H
TP129447	TGCAGTC	chr7B	674,309,026	H	H	H	B	-	B?	H	A	-	A	A	A	A	H	H
TP132	TGCAGAA	chr7B	676,286,093	H	H	H	B	A	H	H	A	A	A	A	A	A	H	H
TP5845	TGCAGAA	chr7B	680,278,011	H	H	H	B	A	H	H	A	-	A	A	A	H	H	H
TP27863	TGCAGCA	chr7B	700,568,686	H	H	H	B	A	H	H	A	-	A	A	A	H	H	A
TP34941	TGCAGCA	chr7B	702,560,697	H	H	H	B	-	H	H	A	-	A	A	A	-	H	-
TP41672	TGCAGCA	chr7B	702,687,580	H	H	H	B	A	A	H	A	A	A	A	A	H	H	A

Nine F_3 plants (WF37, WF92, WF112, WF121, WF149, WF180, WF186, WF189, and WF203) were selected to generate their respective populations for further testing.

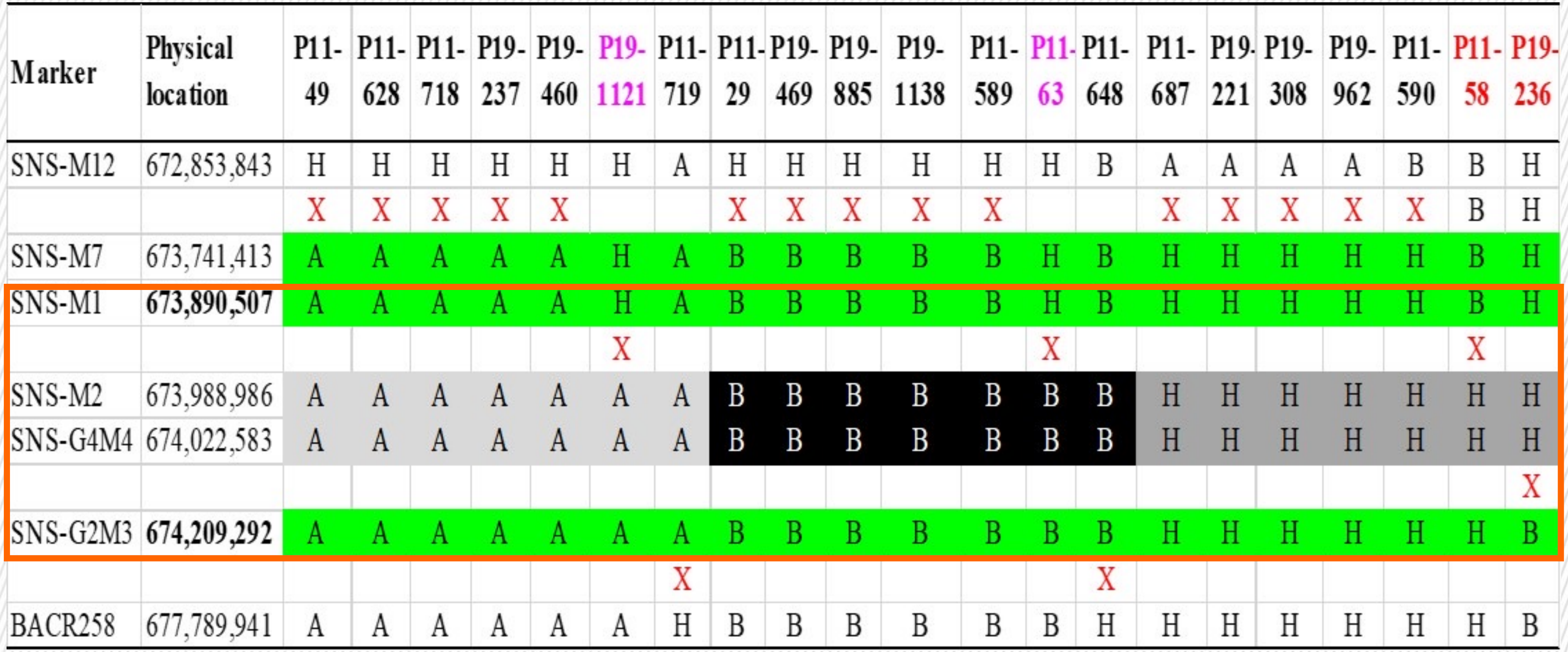
Effects of *QSns.osu-7B* on the number of spikelets per spike

Results

- WF112 F_4 population showed **3:1** segregation ratio for the number of spikelets per spike.
- L20 carried the dominant allele, whereas Y18 carried the recessive allele.
- 1,857 individual plants were screened for crossovers.



2.4. Cloning and validation of *TaCol-B5*

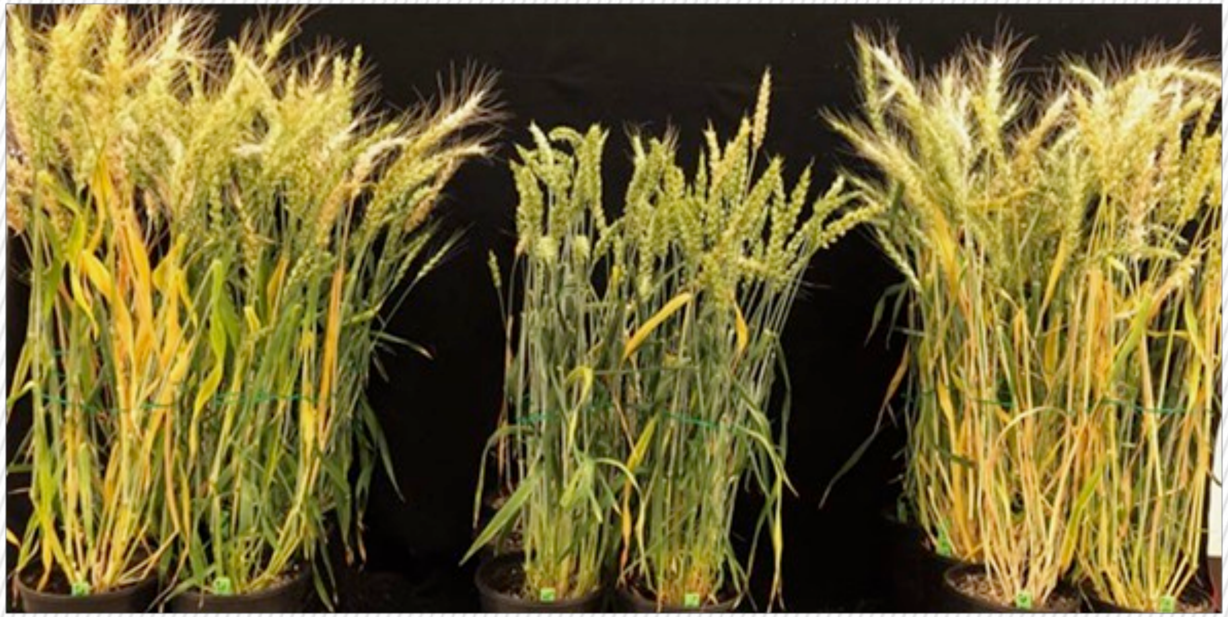


Marker	Physical location	P11-49	P11-628	P11-718	P19-237	P19-460	P19-1121	P11-719	P11-29	P19-469	P19-885	P19-1138	P11-589	P11-63	P11-648	P11-687	P19-221	P19-308	P19-962	P11-590	P11-58	P19-236
SNS-M12	672,853,843	H	H	H	H	H	H	A	H	H	H	H	H	H	B	A	A	A	A	B	B	H
		X	X	X	X	X			X	X	X	X	X			X	X	X	X	X	B	H
SNS-M7	673,741,413	A	A	A	A	A	H	A	B	B	B	B	B	H	B	H	H	H	H	H	B	H
SNS-M1	**673,890,507**	A	A	A	A	A	H	A	B	B	B	B	B	H	B	H	H	H	H	H	B	H
							X							X							X	
SNS-M2	673,988,986	A	A	A	A	A	A	A	B	B	B	B	B	B	B	H	H	H	H	H	H	H
SNS-G4M4	674,022,583	A	A	A	A	A	A	A	B	B	B	B	B	B	B	H	H	H	H	H	H	H
																						X
SNS-G2M3	**674,209,292**	A	A	A	A	A	A	A	B	B	B	B	B	B	B	H	H	H	H	H	H	B
								X							X							
BACR258	677,789,941	A	A	A	A	A	A	H	B	B	B	B	B	B	H	H	H	H	H	H	H	B

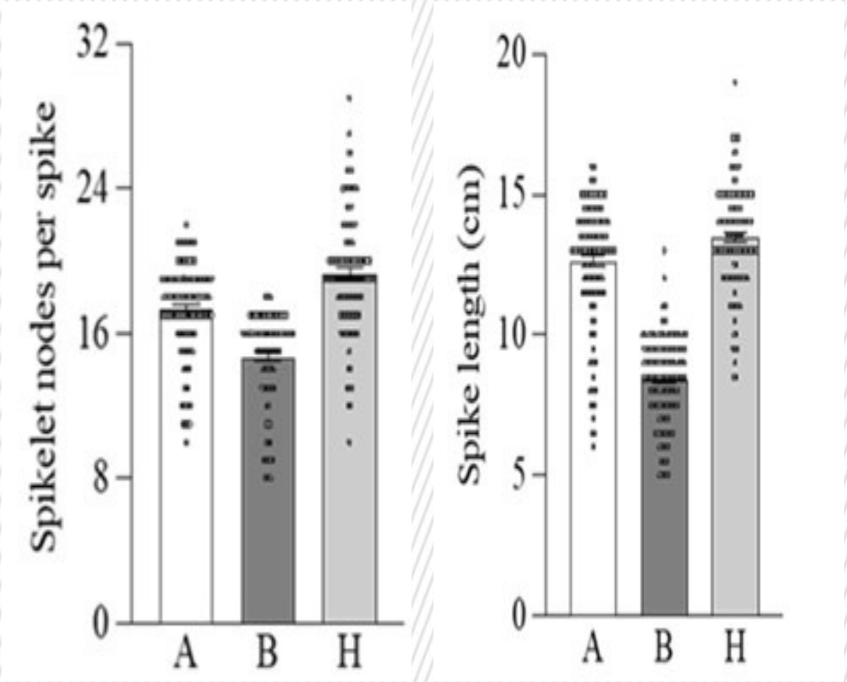
21 crossovers in the targeted region from 1,857 individual $F_{3:4}$ plants.

4 critical crossovers used to determine the candidate region.

Effects of *QSns.osu-7B* in 21 F_5 recombinant plants

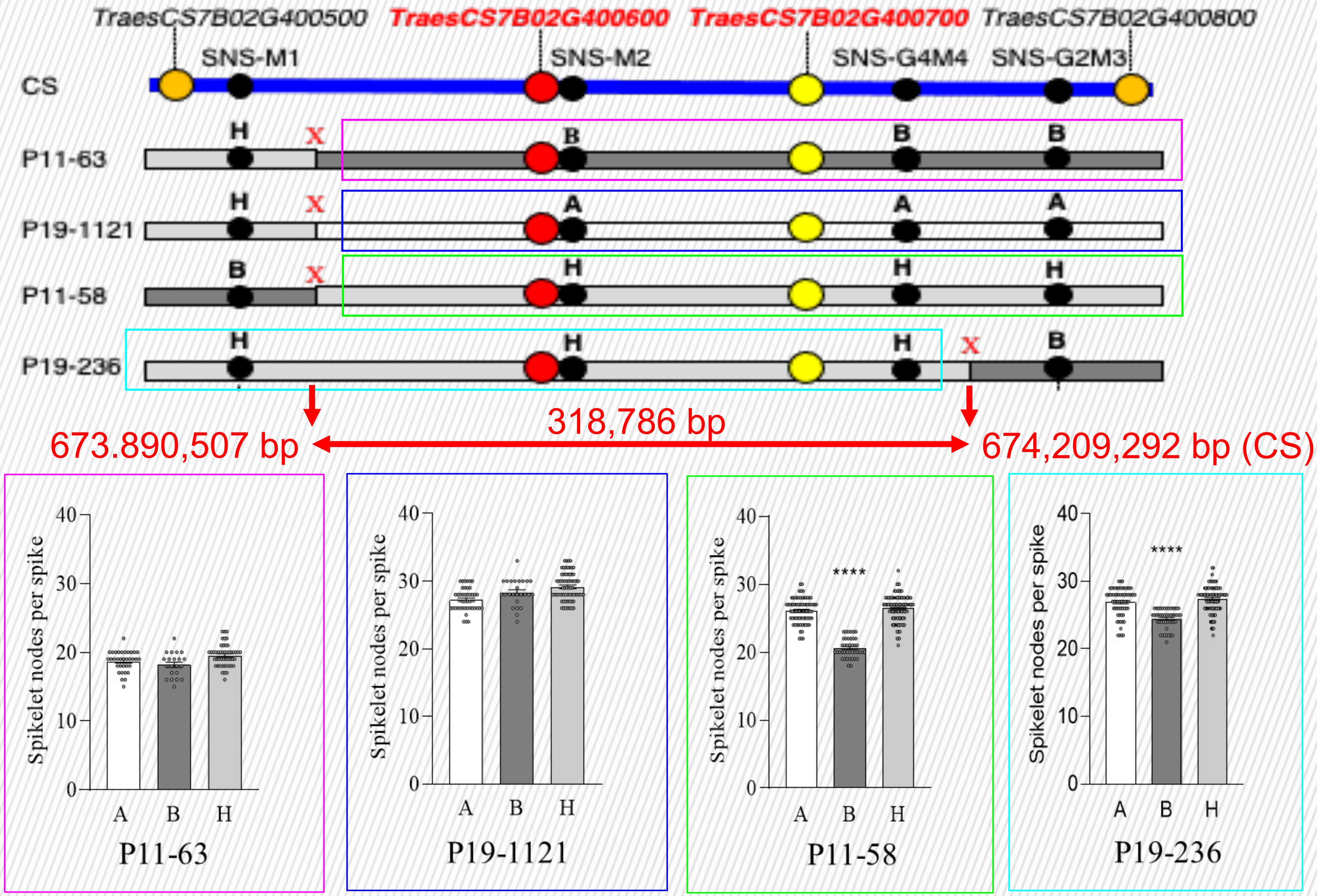


A allele B allele Heterozygotes

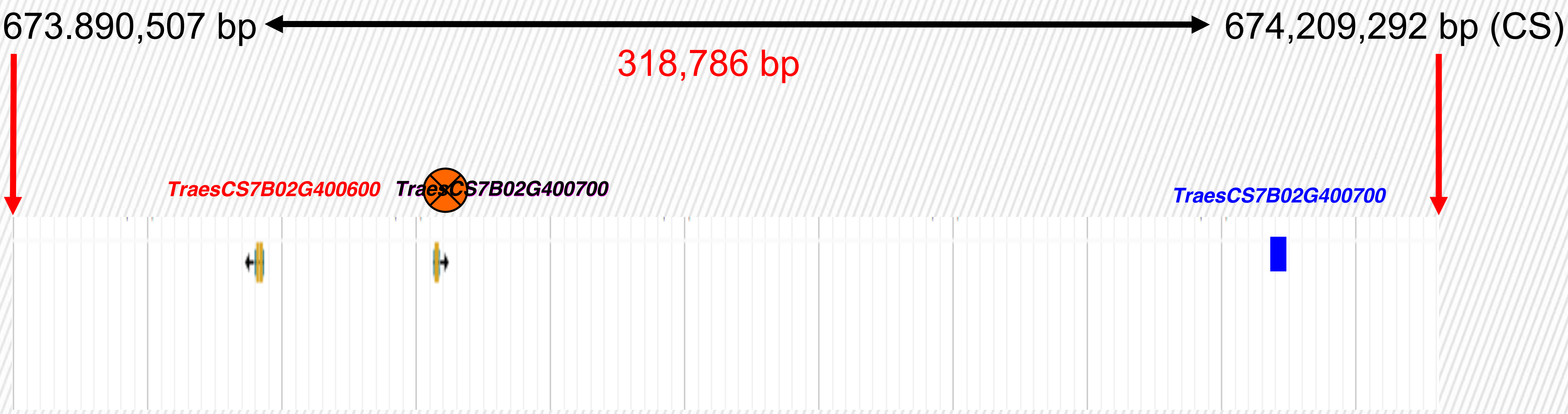


A = L20
B = Y18
H = Het.

Phenotypes and genotypes of 4 critical F_6 recombinant lines



Two candidate genes in the targeted genomic region



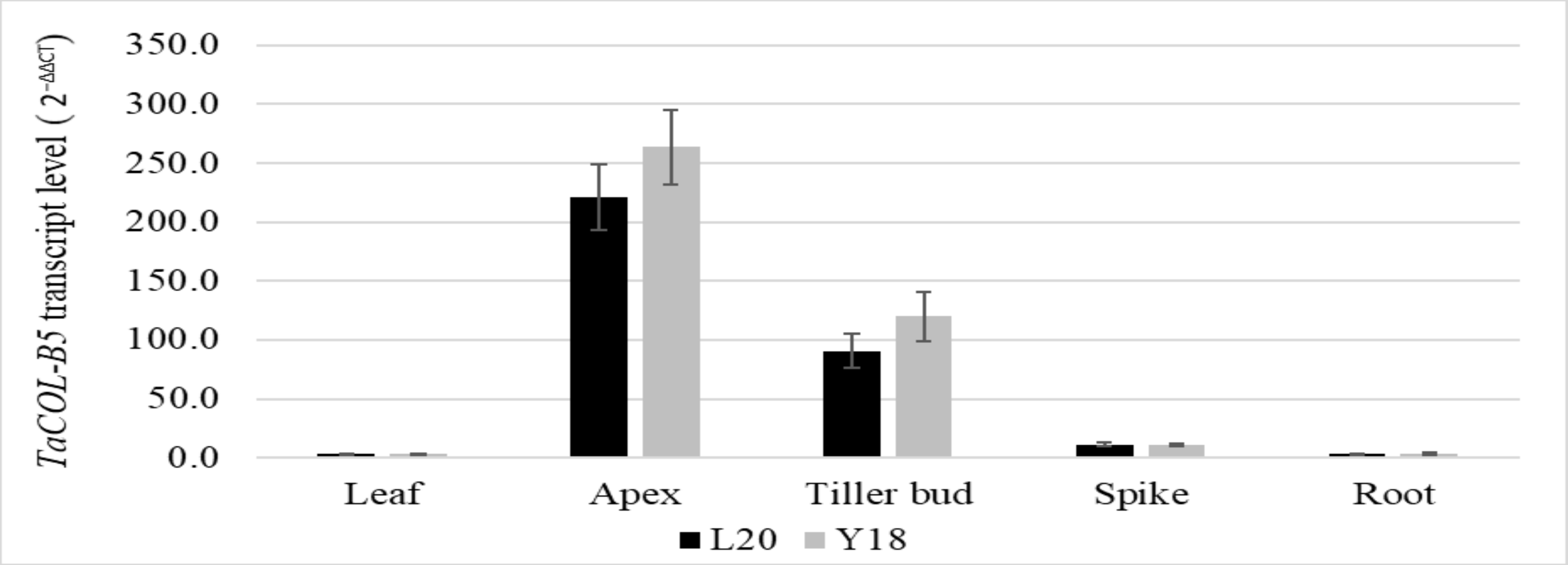
TraesCS7B02G400600 encodes CONSTANS-like protein

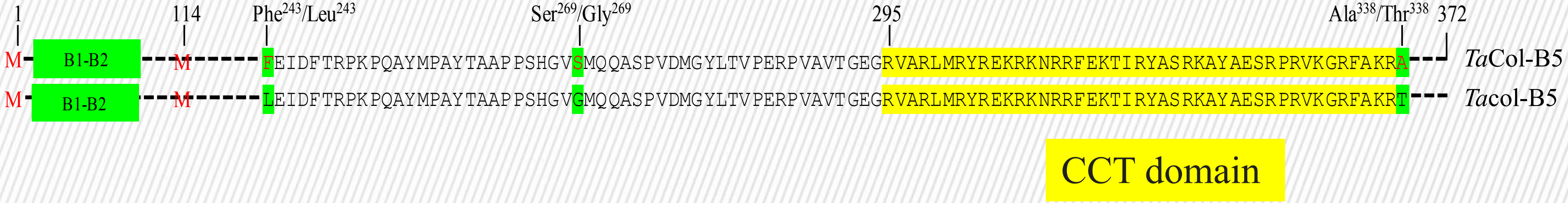
TraesCS7B02G400700 encodes F-box domain containing protein

TraesCS7B02G400700 encodes F-box domain containing protein

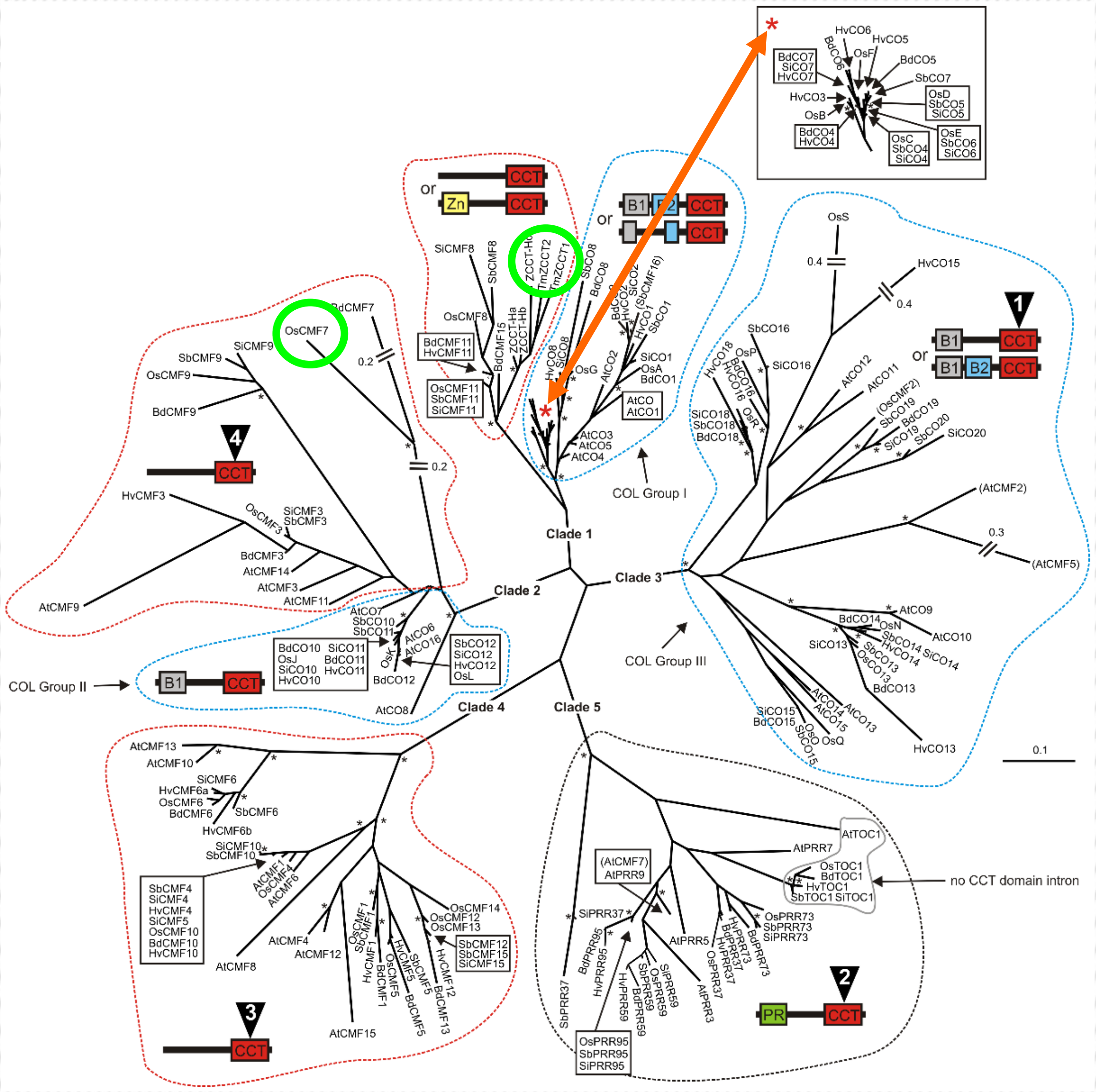
TraesCS7B02G400700: no difference between two alleles

Natural variation in *TraesCS7B02G400600*





Phylogenetic analysis of CMF, COL, PRR and ZCCT proteins, based on the CCT domain



CCT: Constans, Constans-like, TOC1

Constans: A mutant flowers in a ‘constant’ manner regardless of photoperiod.

TOC: Timing of CAB expression

CAB: Chlorophyll a/b binding protein.

CMF: CCT Motif Family

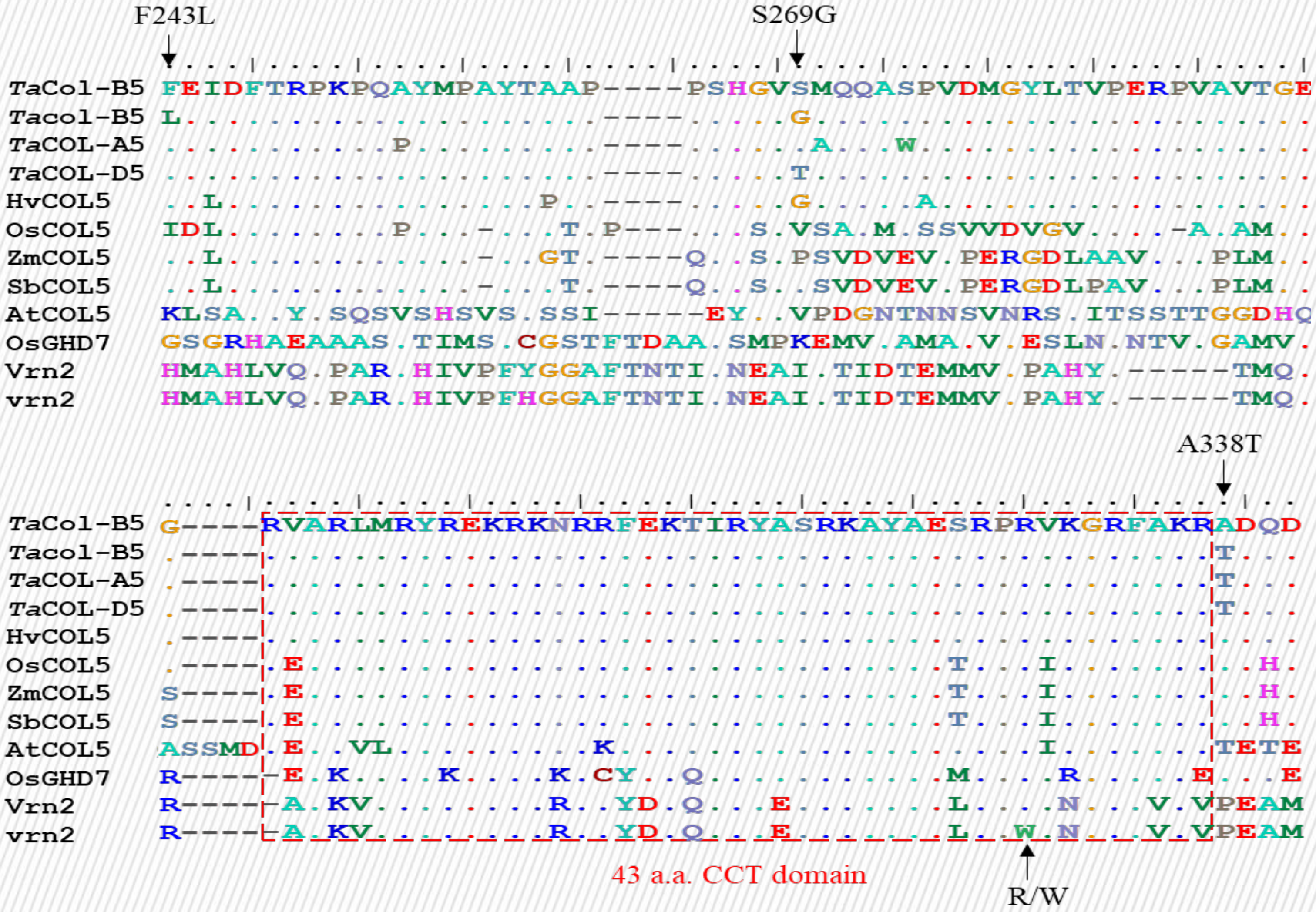
COL: Constans-Like

PRR: Pseudo Response Regulator

ZCCT: Zinc CCT

Cockram *et al*., PLoS ONE 7: e45307 (2012)

Multiple sequence alignment of *Ta*COL-B5 and related proteins



```
          F243L                         S269G
          ↓                             ↓
          ....|....|....|....|....|....|....|....|....|....|....|
TaCol-B5  FEIDFTRPKPQAYMPAYTAAP----PSHGVSMQQASPVDMGYLTVPERPVAVTGE
Tacol-B5  L....................----.....G........................
TaCOL-A5  ...........P.........----......A...W...................
TaCOL-D5  .....................----.....T........................
HvCOL5    ..L...............P..----.....G.....A..................
OsCOL5    IDL........P...-...T.P---...S.VSA.M.SSVVDVGV....-A.AM..
ZmCOL5    ..L............-..GT.----Q..S.PSVDVEV.PERGDLAAV...PLM..
SbCOL5    ..L............-...T.----Q..S..SVDVEV.PERGDLPAV...PLM..
AtCOL5    KLSA..Y.SQSVSHSVS.SSI-----EY..VPDGNTNNSVNRS.ITSSTTGGDHQ
OsGHD7    GSGRHAEAAAS.TIMS.CGSTFTDAA.SMPKEMV.AMA.V.ESLN.NTV.GAMV.
Vrn2      HMAHLVQ.PAR.HIVPFYGGAFTNTI.NEAI.TIDTEMMV.PAHY.-----TMQ.
vrn2      HMAHLVQ.PAR.HIVPFHGGAFTNTI.NEAI.TIDTEMMV.PAHY.-----TMQ.

                                                        A338T
                                                          ↓
          ....|....|....|....|....|....|....|....|....|....|..
TaCol-B5  G----RVARLMRYREKRKNRRFEKTIRYASRKAYAESRPRVKGRFAKRADQD
Tacol-B5  .----...........................................T...
TaCOL-A5  .----...........................................T...
TaCOL-D5  .----...........................................T...
HvCOL5    .----...............................................
OsCOL5    .----.E.............................T...I.........H.
ZmCOL5    S----.E.............................T...I.........H.
SbCOL5    S----.E.............................T...I.........H.
AtCOL5    ASSMD.E..VL.........K...................I.......TETE
OsGHD7    R-----E.K....K....K.CY..Q...........M....R.....E...E
Vrn2      R-----A.KV........R..YD.Q...E.......L....N...V.VPEAM
vrn2      R-----A.KV........R..YD.Q...E.......L..W.N...V.VPEAM
                                                 ↑
```

43 a.a. CCT domain

R/W

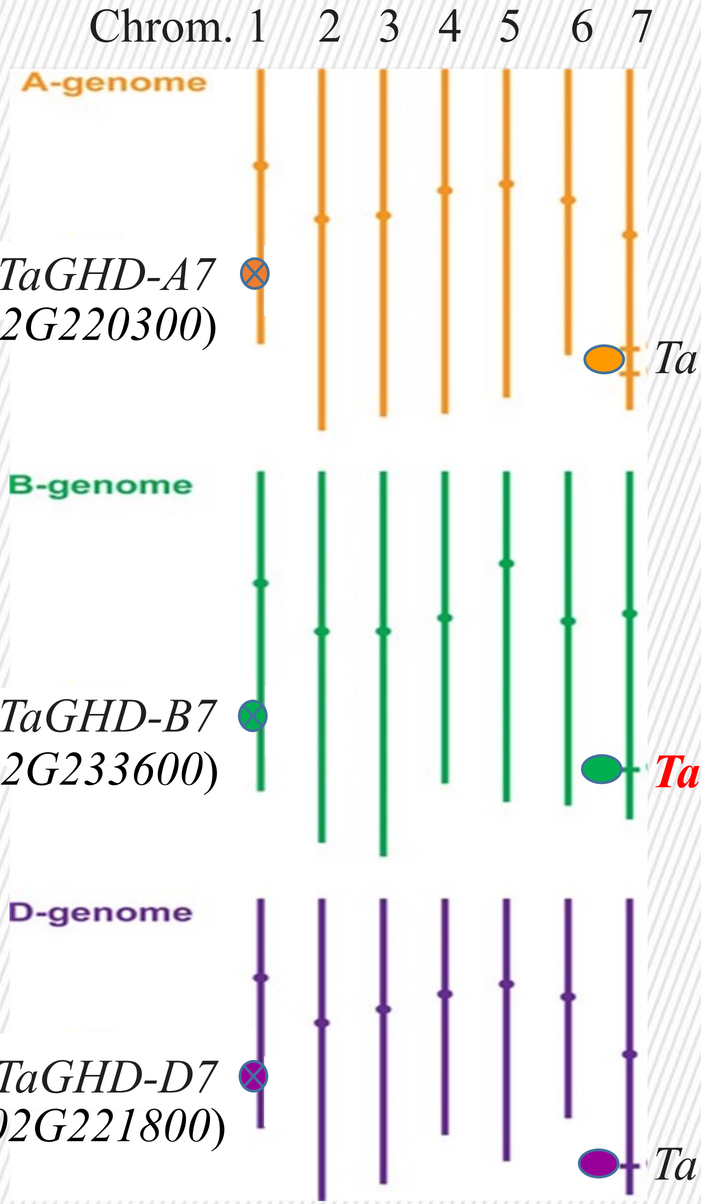
VRN2: Vernalization gene 2, ZCCT1 and ZCCT2.

Yan *et al*., Science. 303: 1640-1644 (2004)

Ghd7: Grains per panicle, plant **h**eight, heading **d**ate.

Xue *et al*., Nat. Genet. 40: 761-767 (2008)

TaCol-B5 was not orthologous to *OsGhd7*



TaGHD-A7
(*TraesCS1A02G220300*)

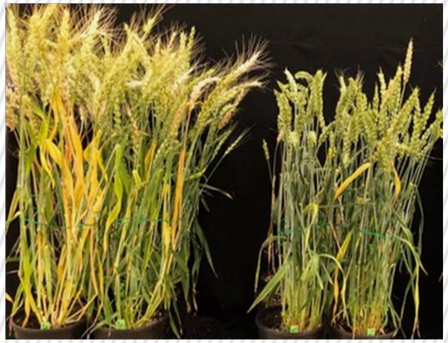
TaCOL-A5 (*TraesCS7A02G497200*)

TaGHD-B7
(*TraesCS1B02G233600*)

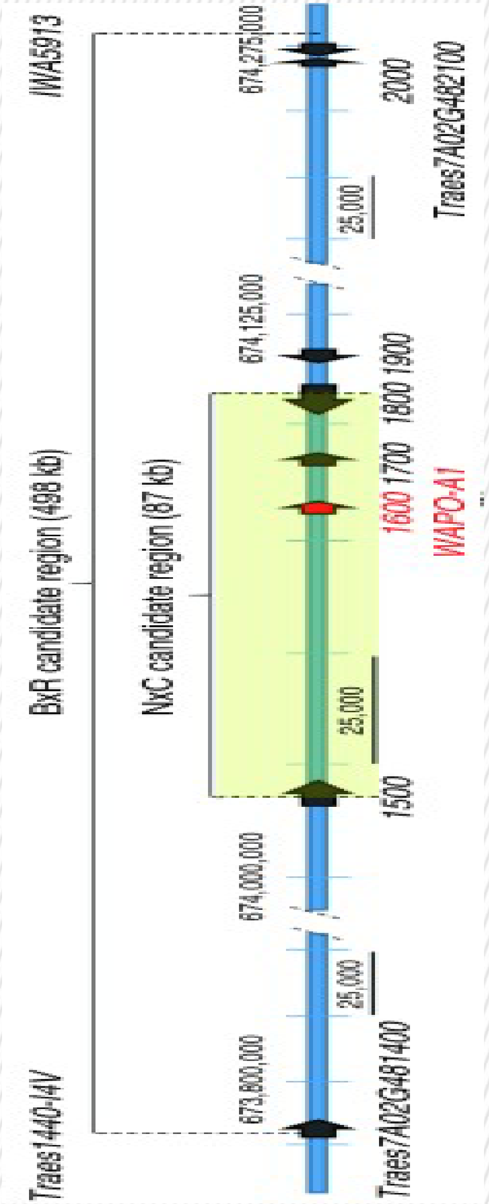
TaCOL-B5 (*TraesCS7B02G400600*)

TaGHD-D7
(*TraesCS1D02G221800*)

TaCOL-D5 (*TraesCS7D02G484400*)



TaCOL5 is close to *WAPO1* on chromosome 7AL



APO1: Aberrant Panicle Organization 1

TraesCS7A02G481600

186 genes

TaCOL-A5

TraesCS7A02G497200

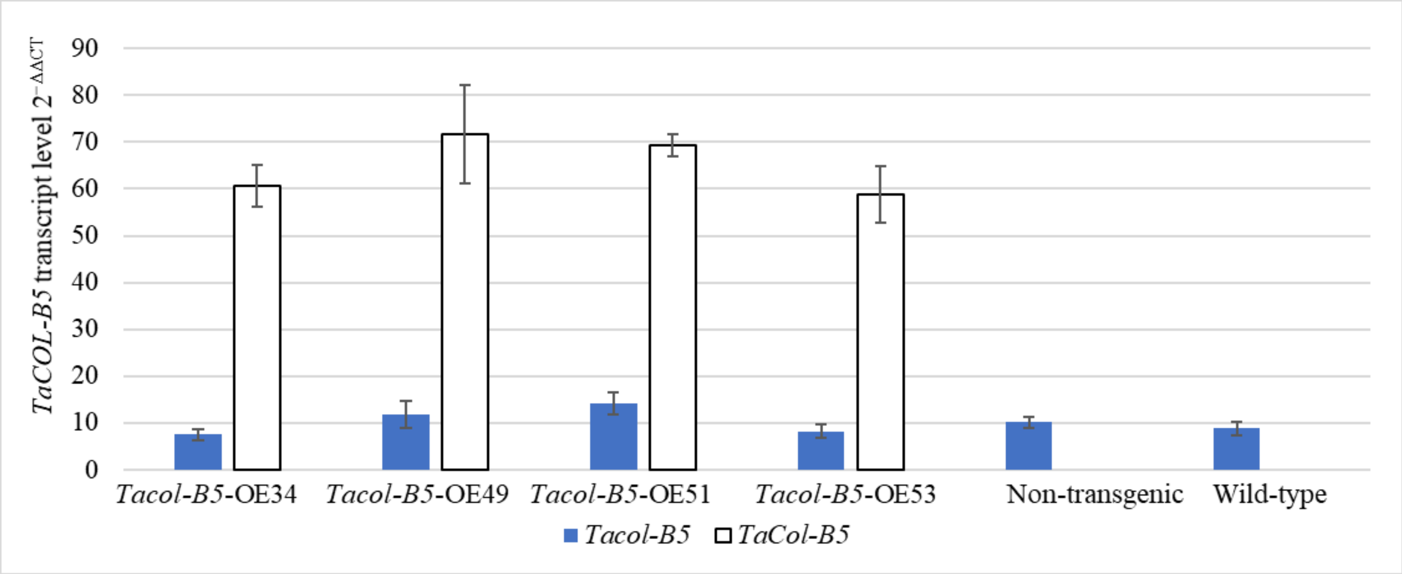
Kuzay et al., Theor Appl Genet. 132: 2689-2705 (2019)

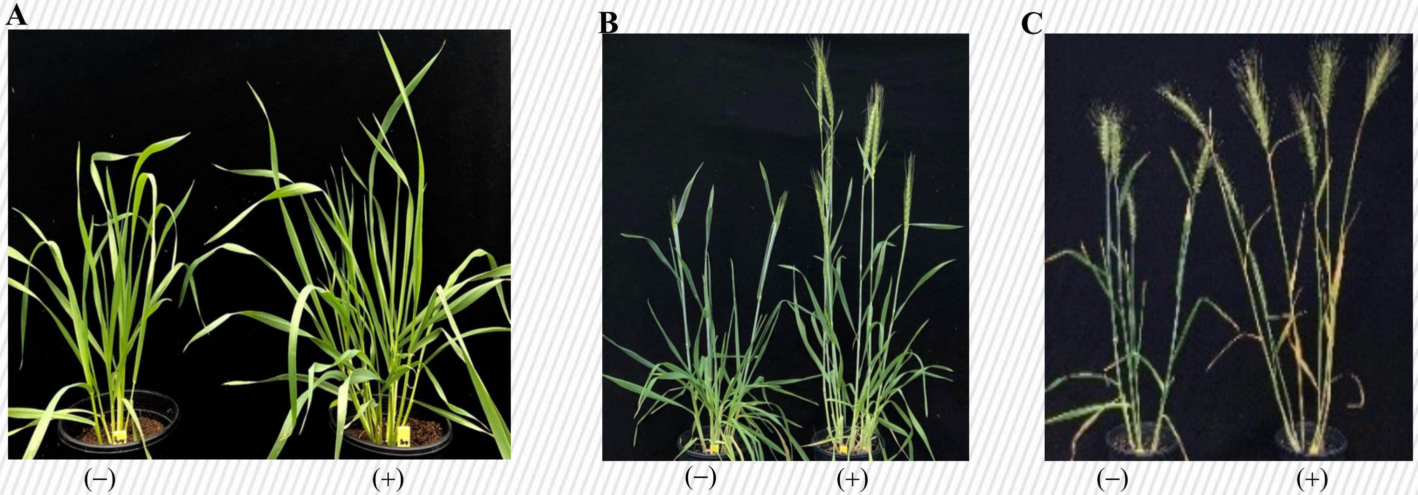
Voss-Fels et al., Theor Appl Genet. 132: 2707–2719 (2019)

Ward et al., PLoS ONE 14, e0208217 (2019)

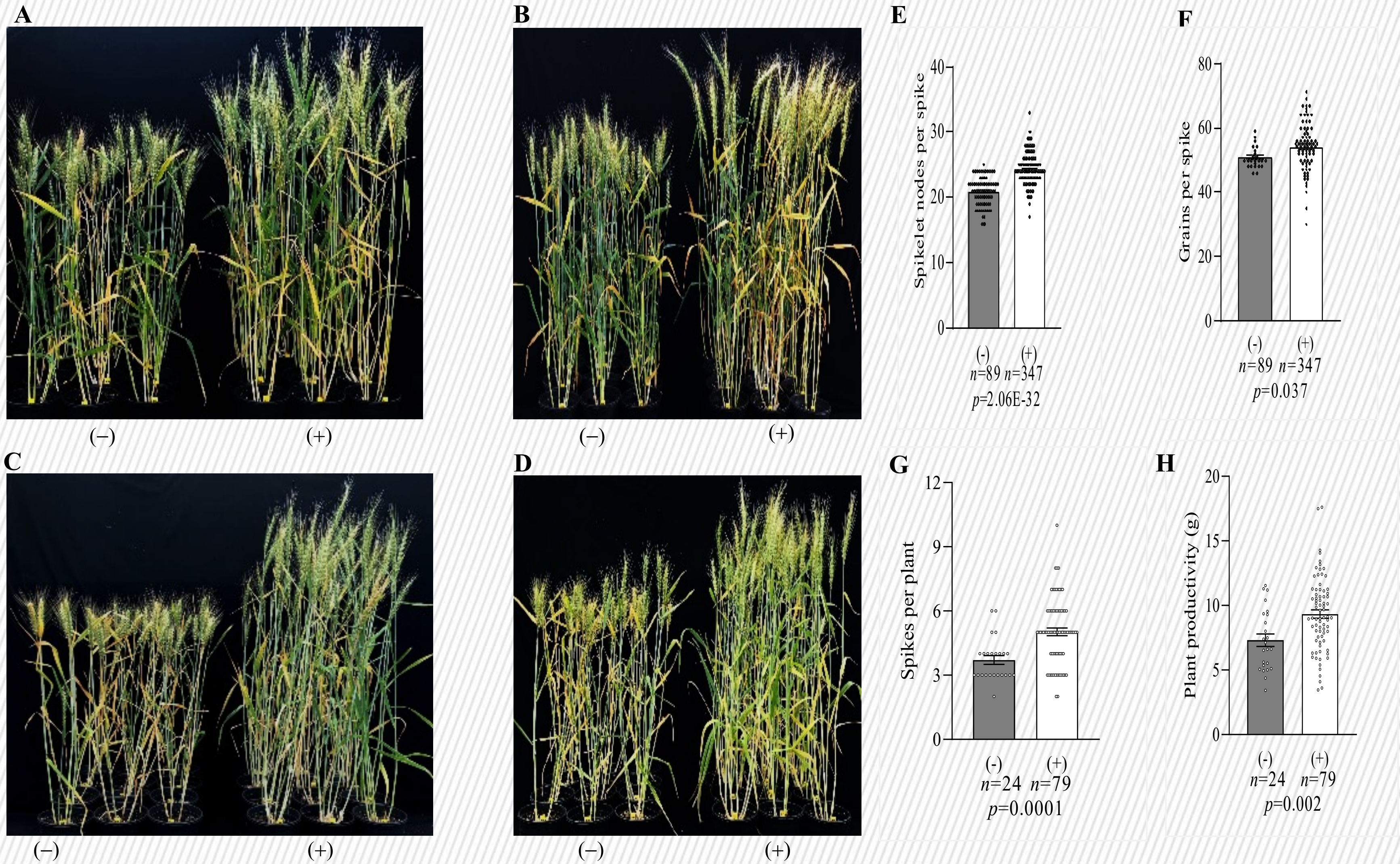
Muqaddasi et al., Sci Rep-Uk. 9: 13853 (2019)

Constitutive overexpression of *TaCol-B5* from CItr 17600 into Yangmai18



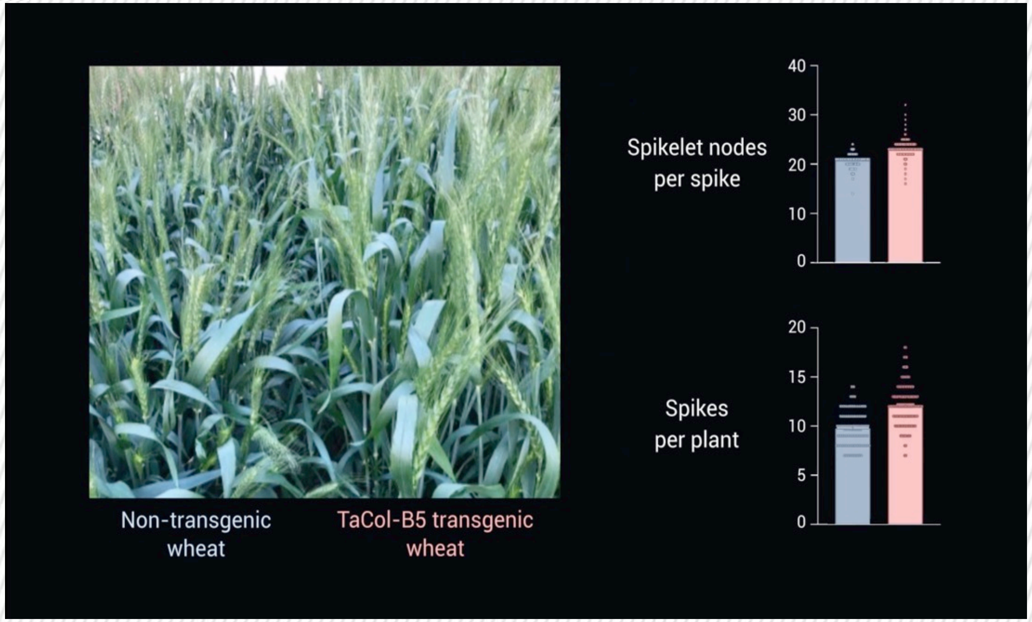


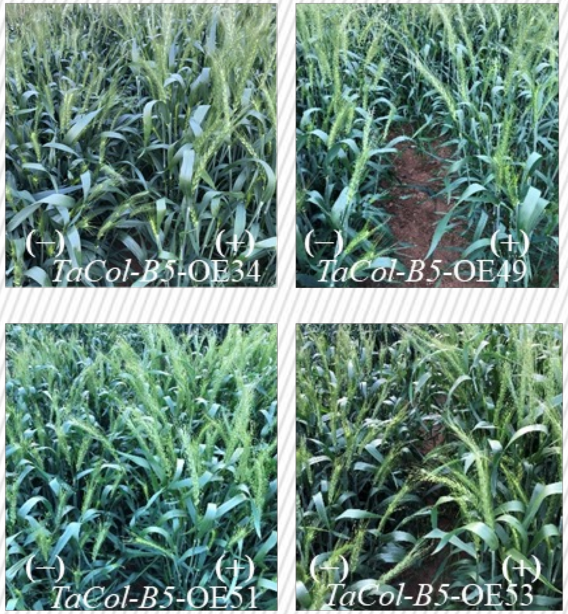
Effects of *TaCol-B5* in transgenic Yangmai18

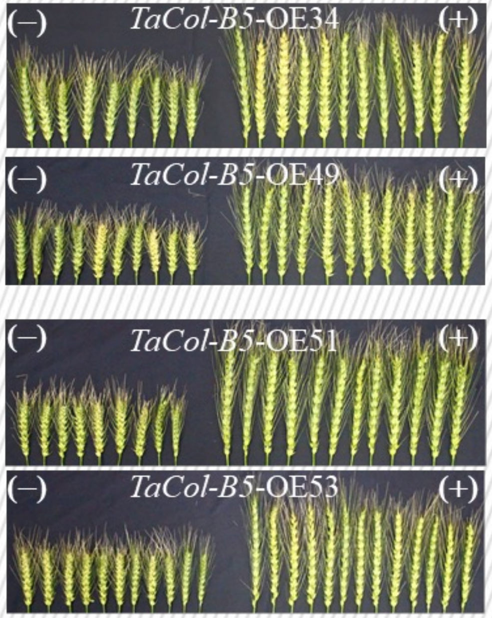


3. Post-cloning activities

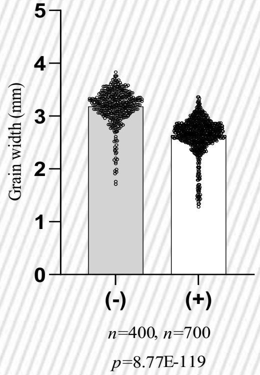
3.1. Validation of T_2 populations in the field

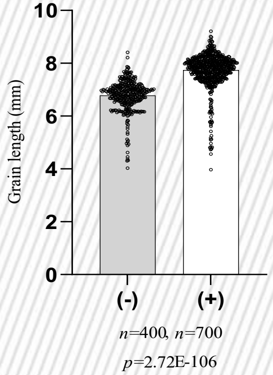




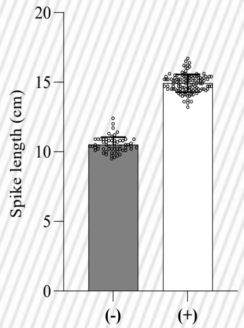


T_2 plants were tested either in four rows (transgenic) or in two rows (non-transgenic), with two replicates and at a relatively low plant density. Rows were 2 m long and spaced 0.25 m apart. Plant density was fixed at 40 plants per m^2.

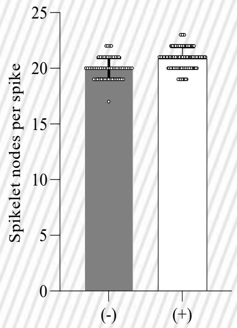




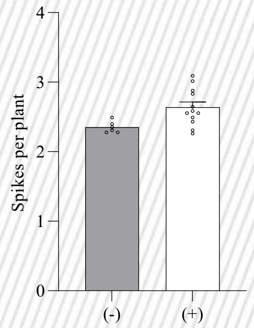
Effects of *TaCol-B5* on field-based grain yield in T_3 populations



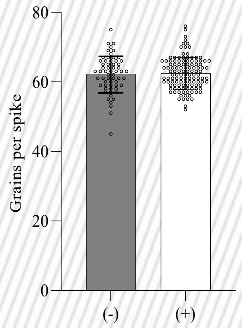
n=60 *n*=120
p=3.5E-100



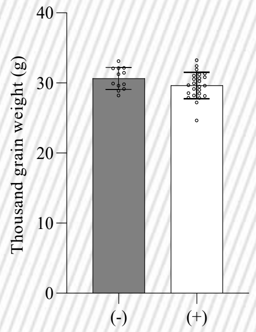
n=60 *n*=120
p=6.44E-08



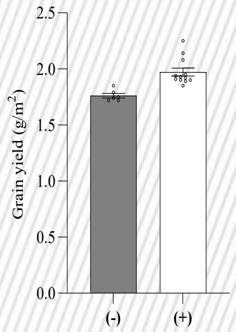
n=6 *n*=12
p=0.022



n=60 *n*=120
p=0.789



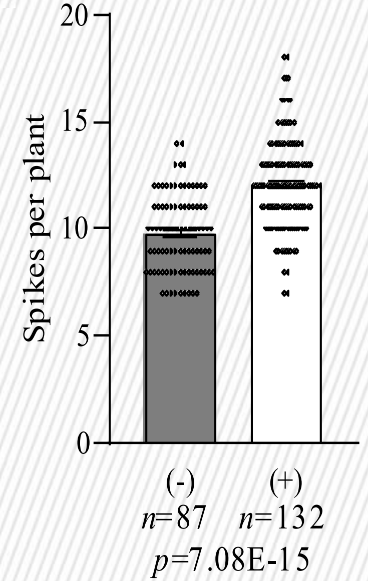
n=12 *n*=24
p=0.121

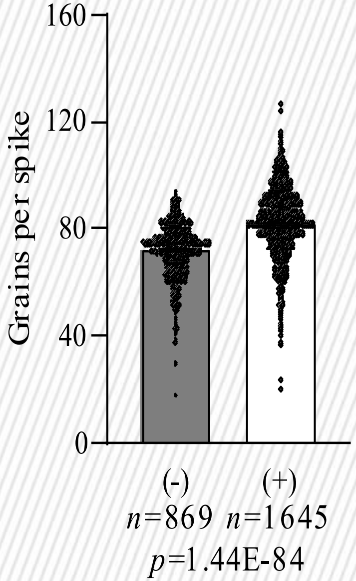


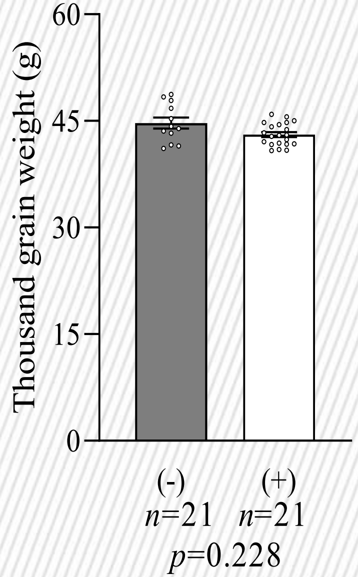
n=6 *n*=12
p=0.0009

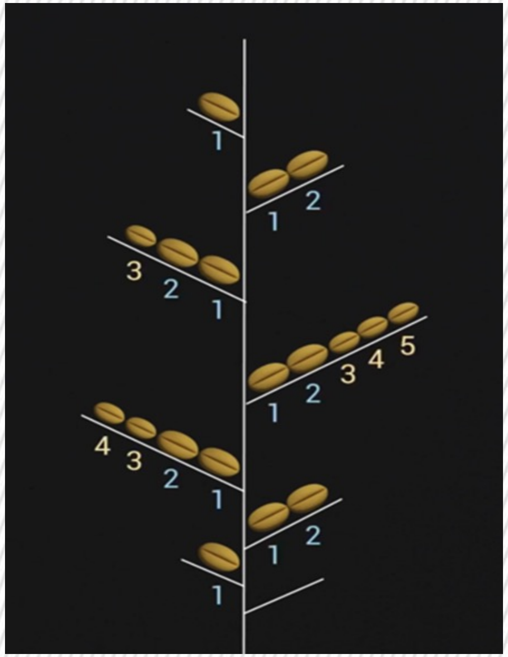
Overall, grain yield of the four transgenic Yangmai18 lines was increased **11.9%** in these 6 m^2 plots, compared to the non-transgenic Yangmai18. For the individual transgenic lines, grain yield increased 7.8% in *Ta*Col-B5-OE34, 19.8% in *Ta*Col-B5-OE49,11.6% in *Ta*Col-B5-OE51, and 8.4% in *Ta*Col-B5-OE53, compared to the non-transgenic plants.

Effects of *TaCol-B5* on trade-off between yield components









Trade-off between spikes per plant and grains per spike

Trade-off between grains per spike and grain weight

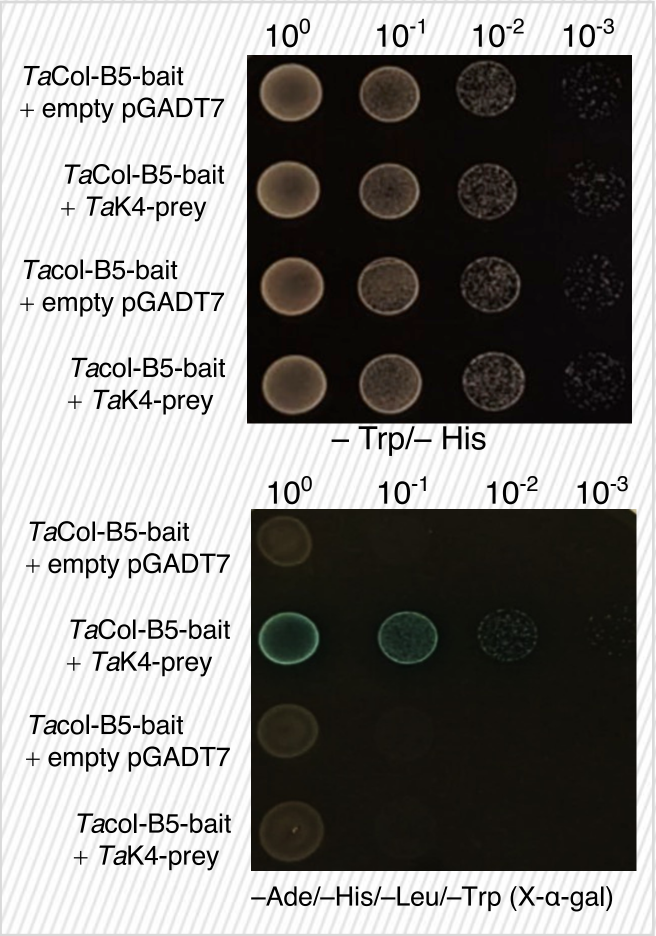
3.2. Mechanisms of the cloned gene underlying the traits

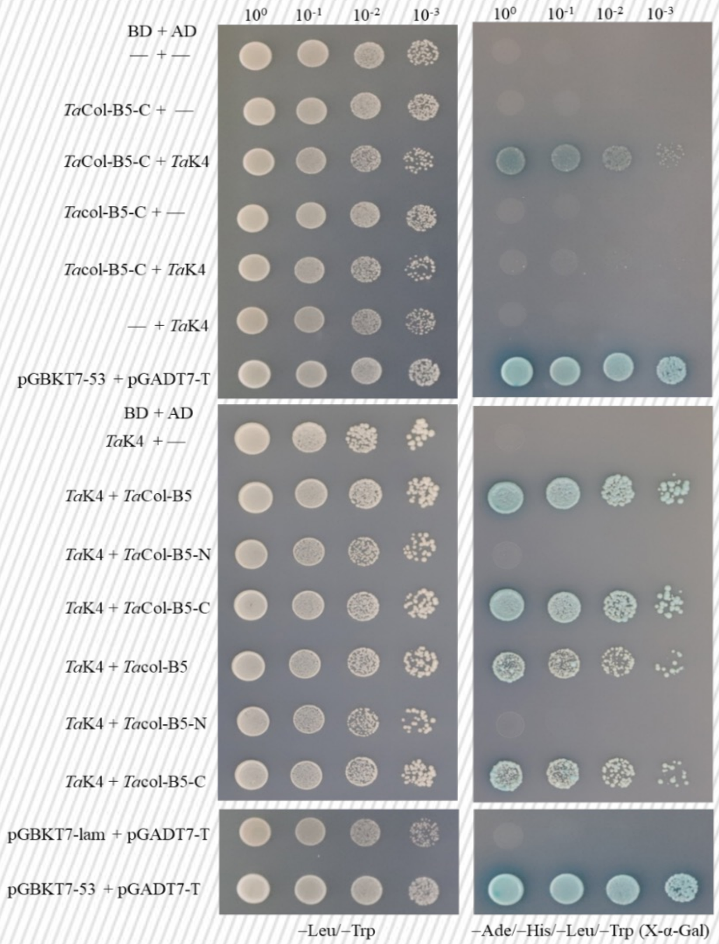
Position 243: Phenylalanine vs. Leucine

Position 269: Serine vs. Glycine

Position 338: Alanine vs. Tyrosine

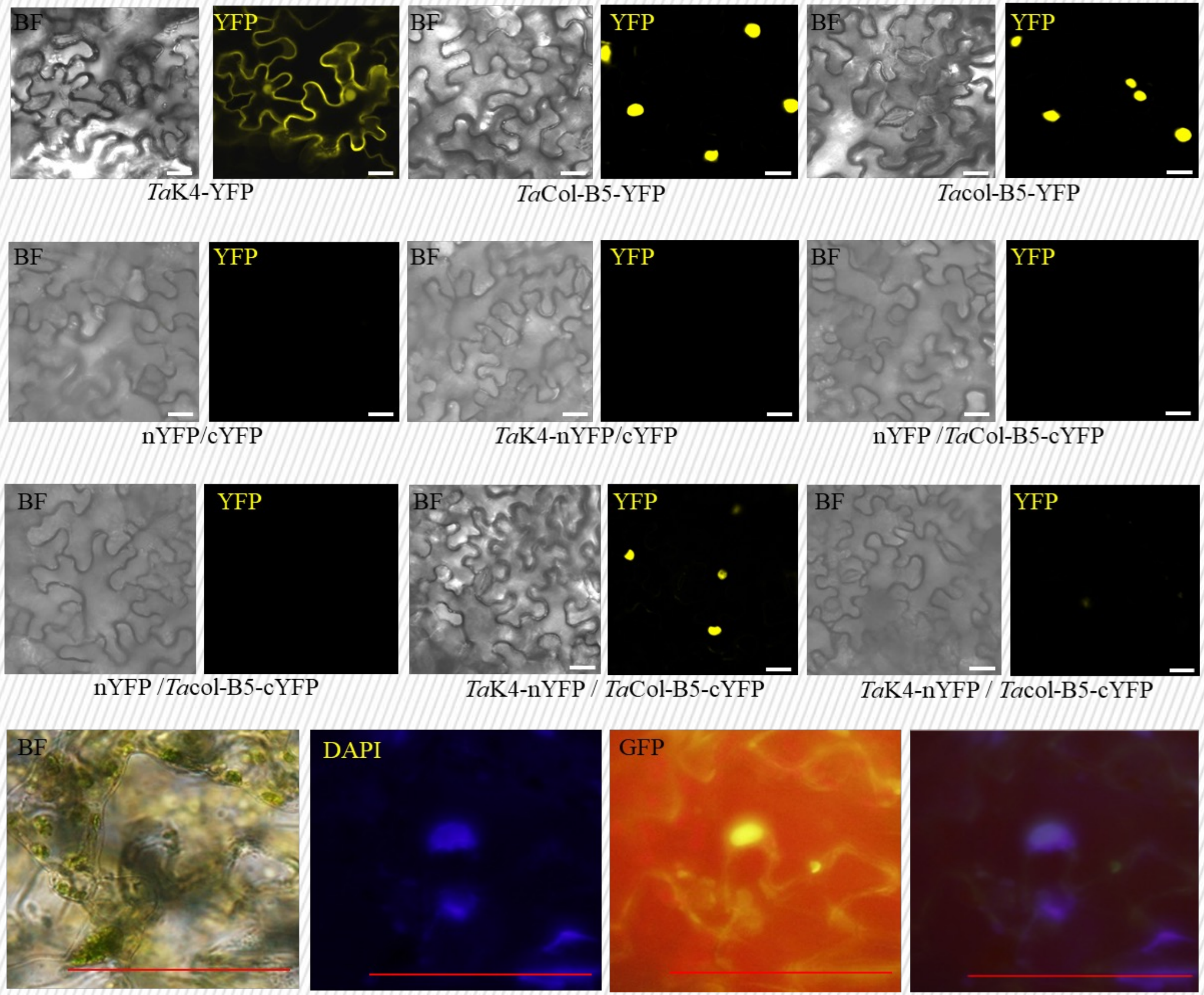
Interactions of *Ta*COL-B5 with *Ta*K4 in yeast





*Ta*K4 is an ortholog of rice *OsK4* encoding a serine/threonine protein kinase.

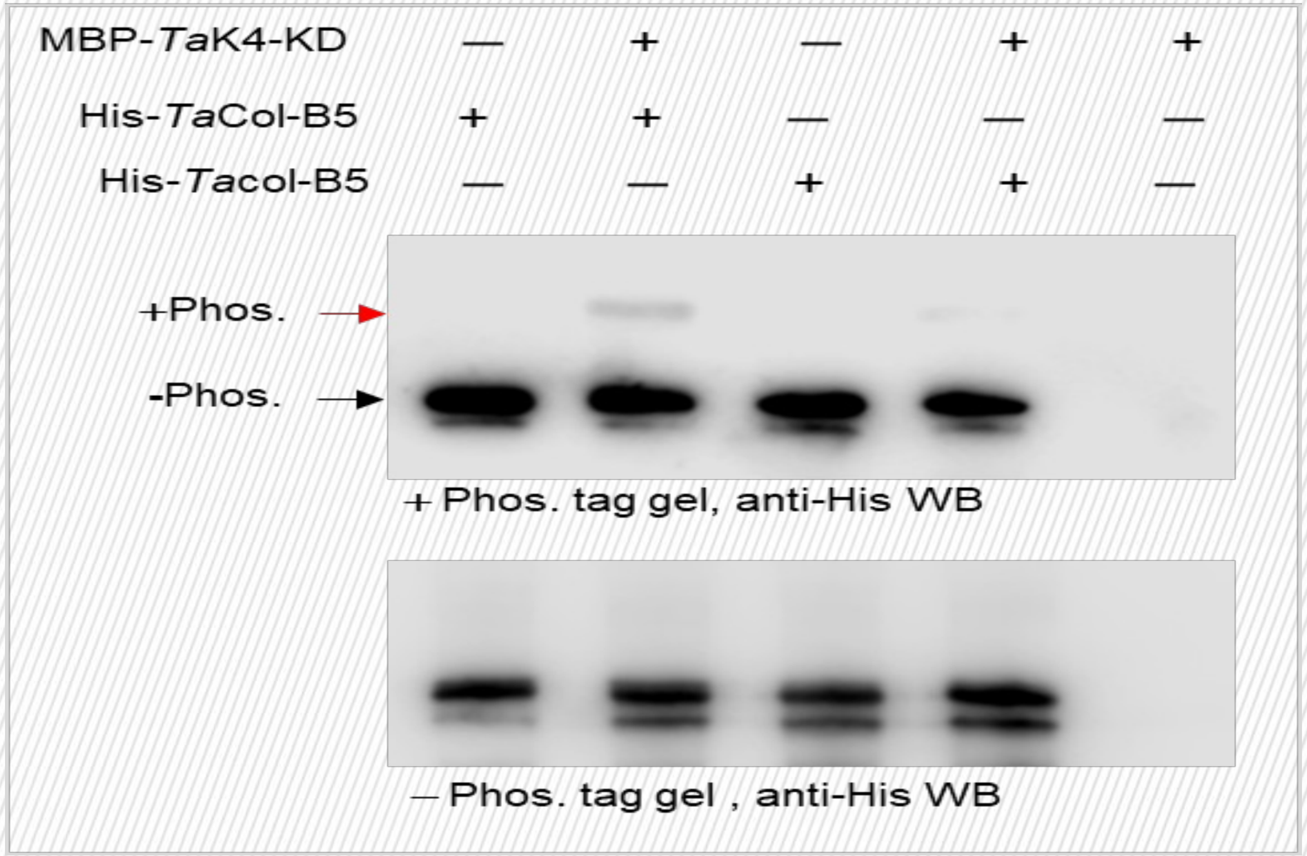
In vivo nteractions of *Ta*COL-B5 with *Ta*K4 in tobacco leaves

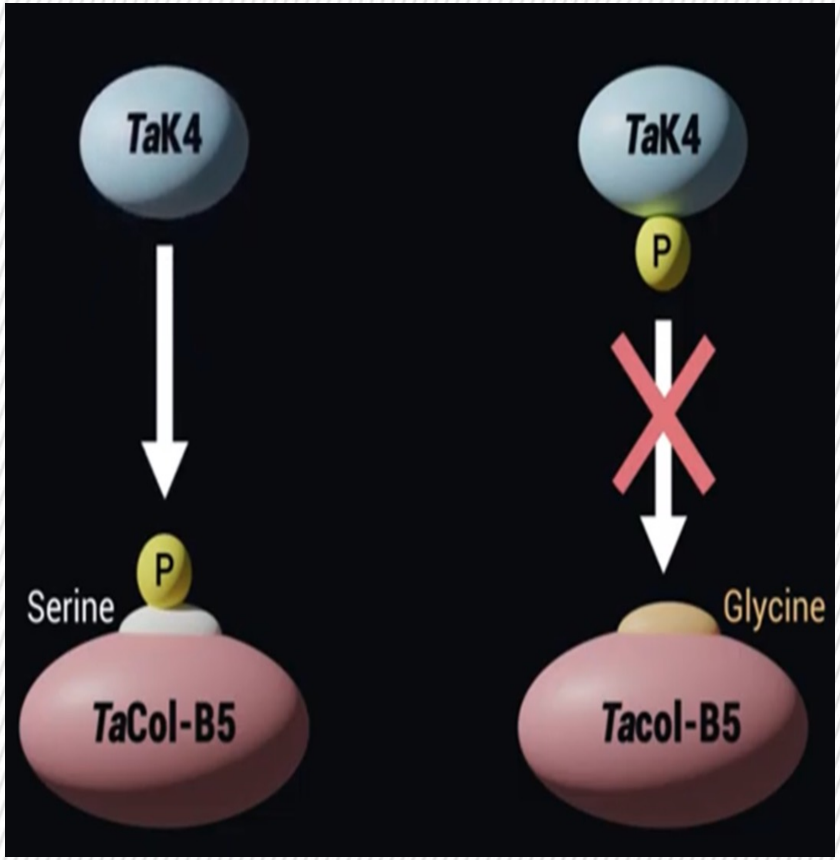


Phosphorylation of *Ta*Col-B5 by *Ta*K4

F130L S269G A338T
*Ta*Col-B5 ---- Phe ------------ Ser ------------ Ala ---- 372 a.a.
*Ta*col-B5 ---- Leu ------------ Gly ------------ Thr ---- 372 a.a.

Three amino acid substitutions between *Ta*Col-B5C and *Ta*col-B5, and two of them that could be differentially phosphorylated, are indicated in red.

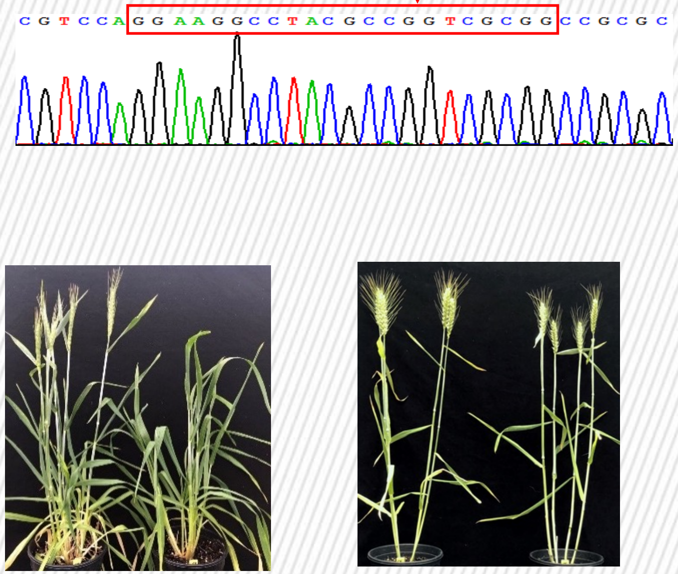




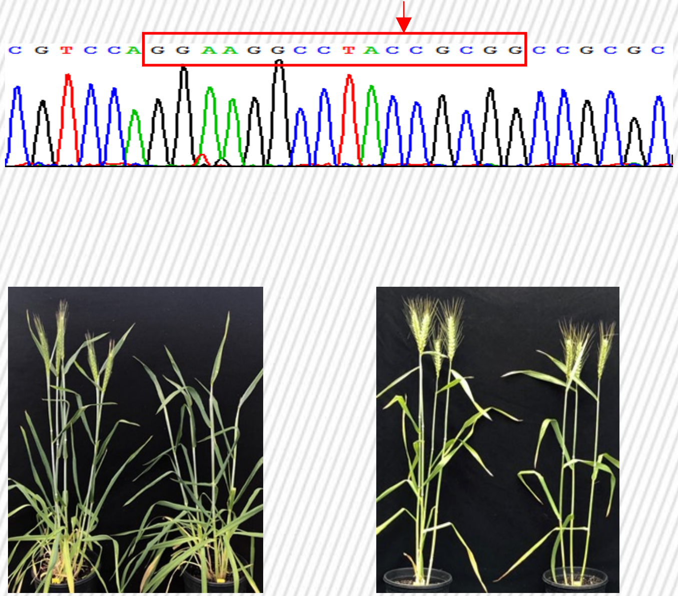
Phosphorylation of *Ta*Col-B5 by *Ta*K4. An *in vitro* kinase assay was performed with purified His-tagged *Ta*COL-B5 proteins.

Pleiotropic effects of edited CCT domain in *Ta*col-B5 protein

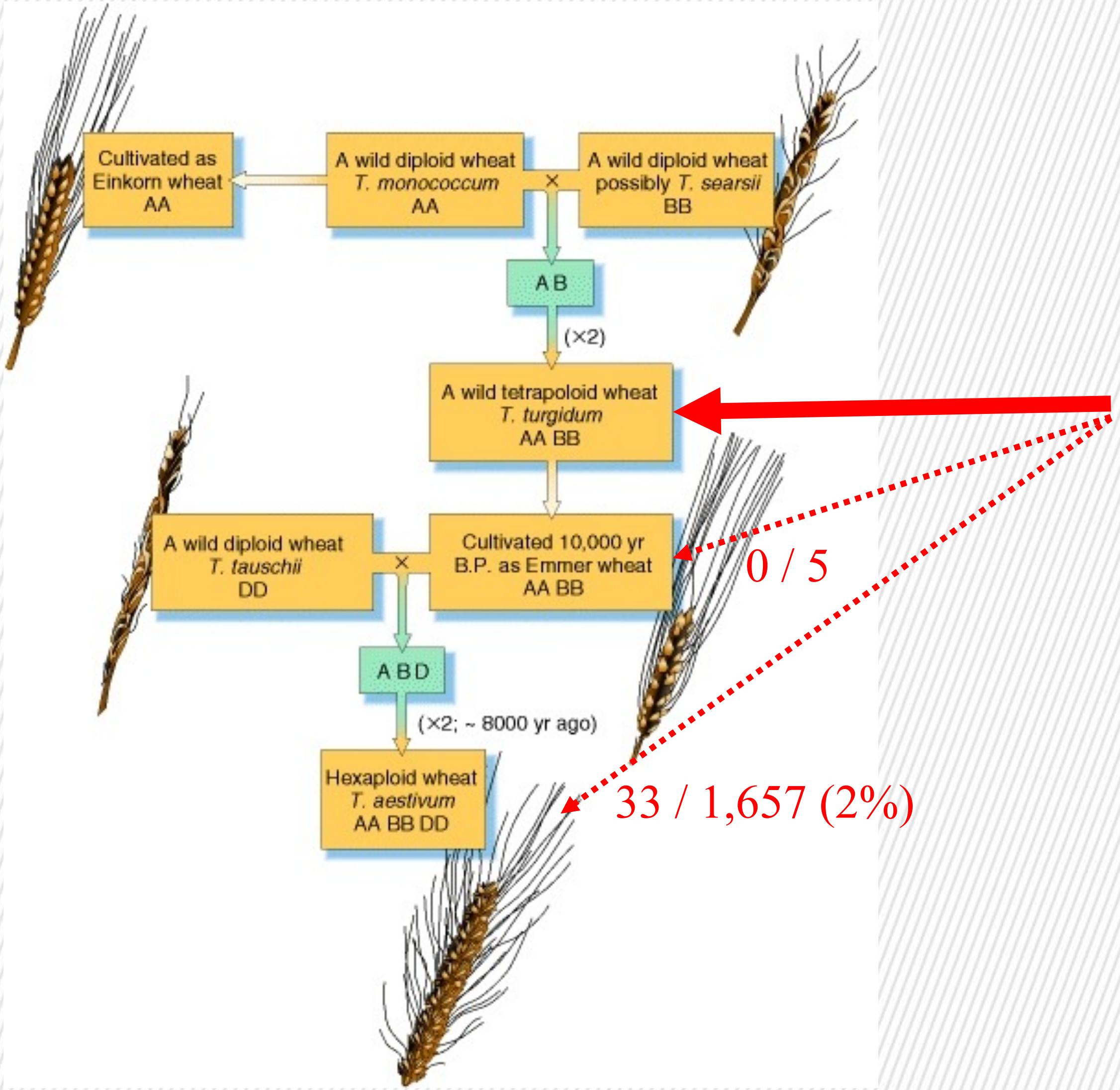
Tacol-B5-ED1 GGAAGGCCTACGCCGAGTCGCGG



Tacol-B5-ED12 GGAAGGCCTACGCCGAGTCGCGG



3.3. Origin and distribution of ***TaCol-B5***



TaCol-B5 in cultivar 'Zavitan'

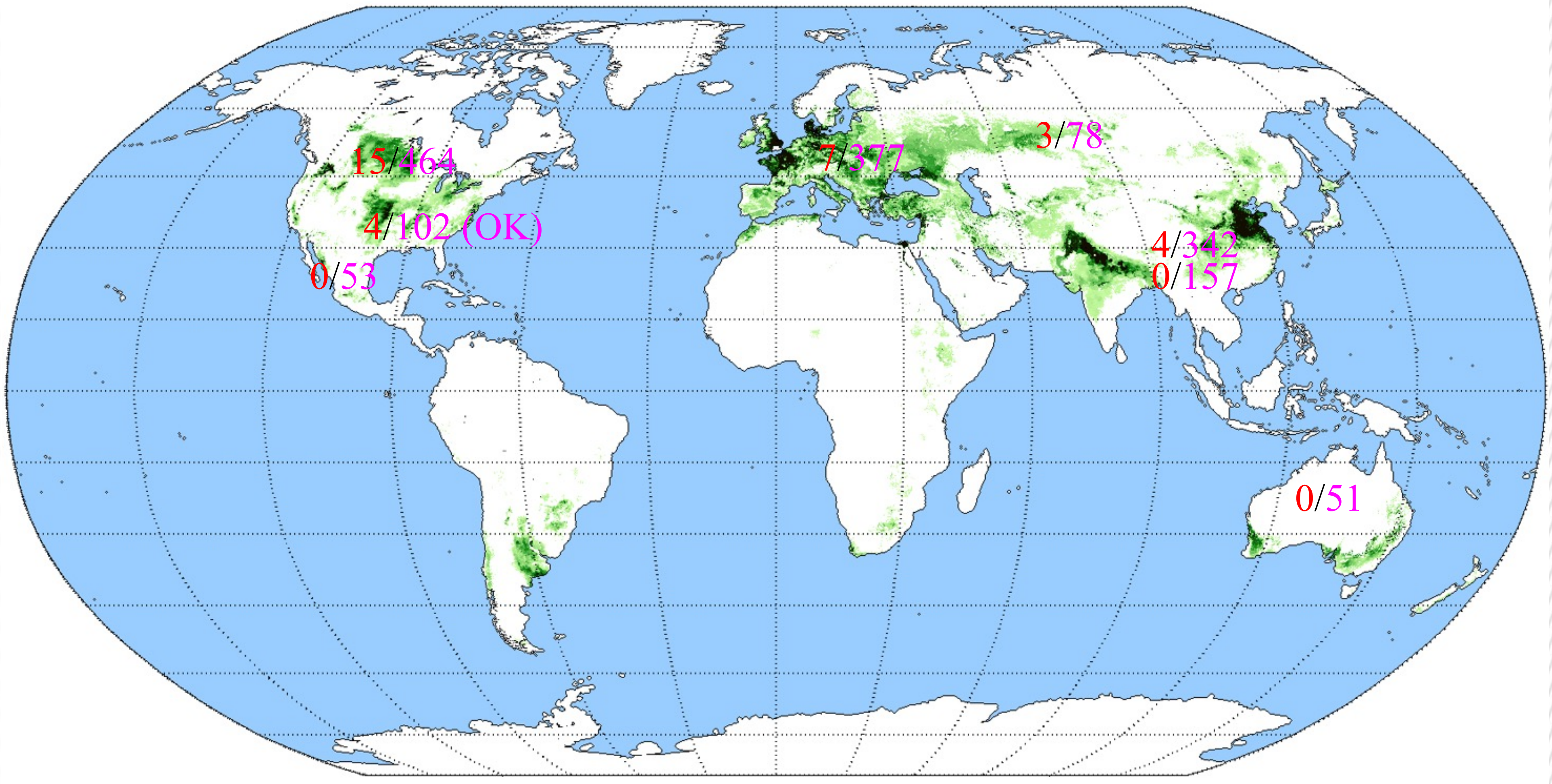
Avni et al., Science, 357, 93-97 (2017)

0 / 5

33 / 1,657 (2%)

https://www.cerealsdb.uk.net/cerealgenomics/WheatBP/Documents/DOC_Evolution.php

The distribution of *TaCol-B5*/*Tacol-B5* allele in global wheat accessions



Summary

- Yield genes can be cloned using the map-based cloning approach
- *TaCol-B5* was cloned using the phenotypes from the controlled conditions but the gene effects were validated in the field
- Different technologies are needed to reveal the functional mechanisms of the cloned genes
- The dominant *TaCol-B5* allele for the higher yield is rare, but it is still accessible from modern wheat cultivars in different continents.

2’21” video on the *TaCol-B5* story
https://www.youtube.com/watch?v=Vy01eOC6hzM

4. Perspectives

Science

PERSPECTIVES

PLANT BIOLOGY

The quest for optimal plant architecture

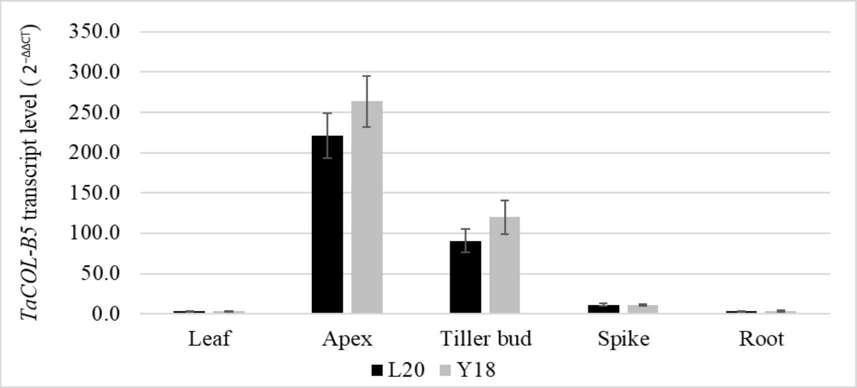
Changes in plant architecture can improve cereal crop yield

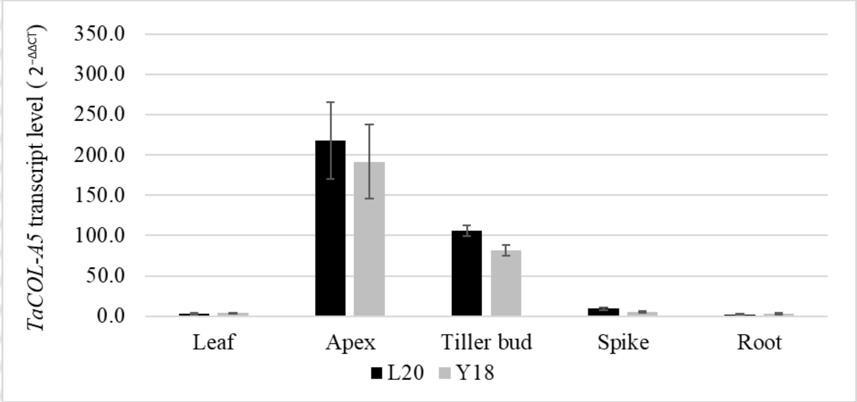
By **G. Wilma van** Esse

https://www.science.org/doi/10.1126/science.abo7429

4. Perspectives

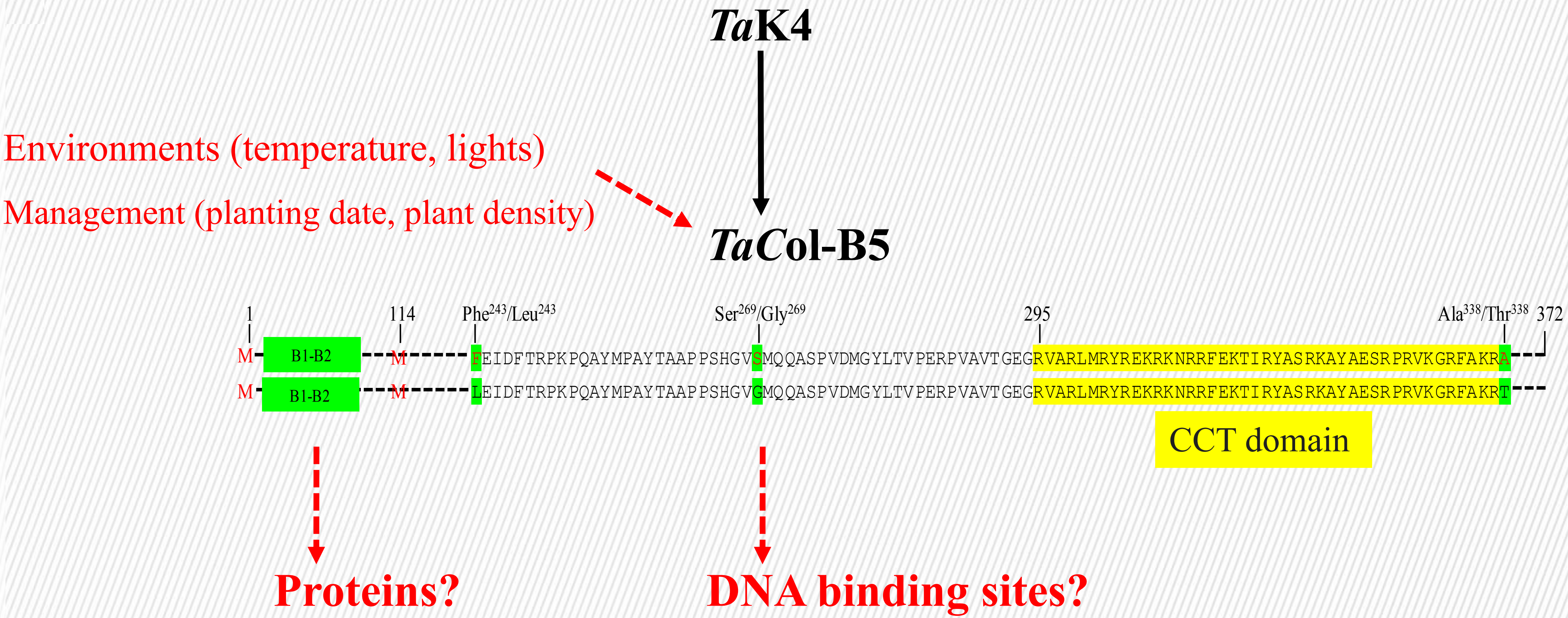
What functions do homoeologous *TaCOL-A5* and *TaCOL-D5* have in wheat?





Perspectives

Grain yield pathways?



***Ta*K4**

Environments (temperature, lights)

Management (planting date, plant density)

***Ta*Col-B5**

Proteins?

DNA binding sites?

Perspectives

Will wheat breeders like to utilize the *TaCol-B5* allele to increase grain yield accompanied by an increase in plant height?

Norman Borlaug, The Noble Prize Laureate

One dream: the super rice grows taller than sorghum, the ears are as long as a broom, and the seeds are the size of peanuts. It is called the dream of enjoying the coolness under Hexia. - Longping Yuan, The father of hybrid rice.

Special acknowledgements

Science

AAAS

ROAD MAP FOR WHEAT

RESEARCH

RESEARCH ARTICLE SUMMARY

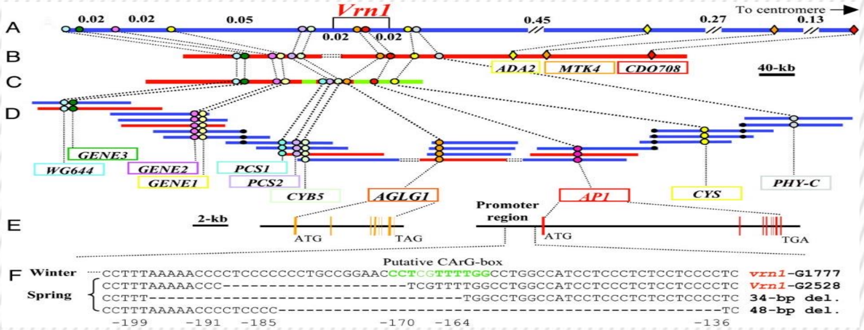
WHEAT GENOME

Shifting the limits in wheat research and breeding using a fully annotated reference genome

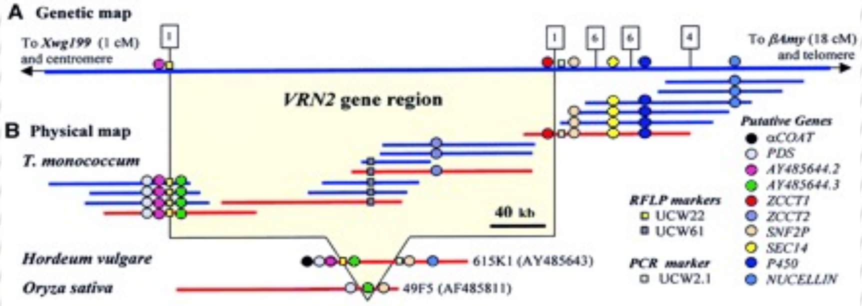
International Wheat Genome Sequencing Consortium (IWGSC)*

Thanks and appreciations to the scientists who sequenced the wheat genome!!!

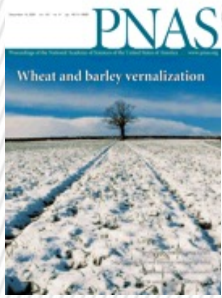
Chromosome walking is no longer needed in gene cloning!!!

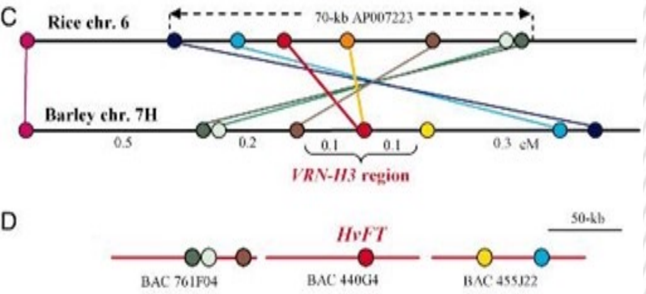


VRN1, **Yan et al. PNAS, 100:6263-6268 (2003)**



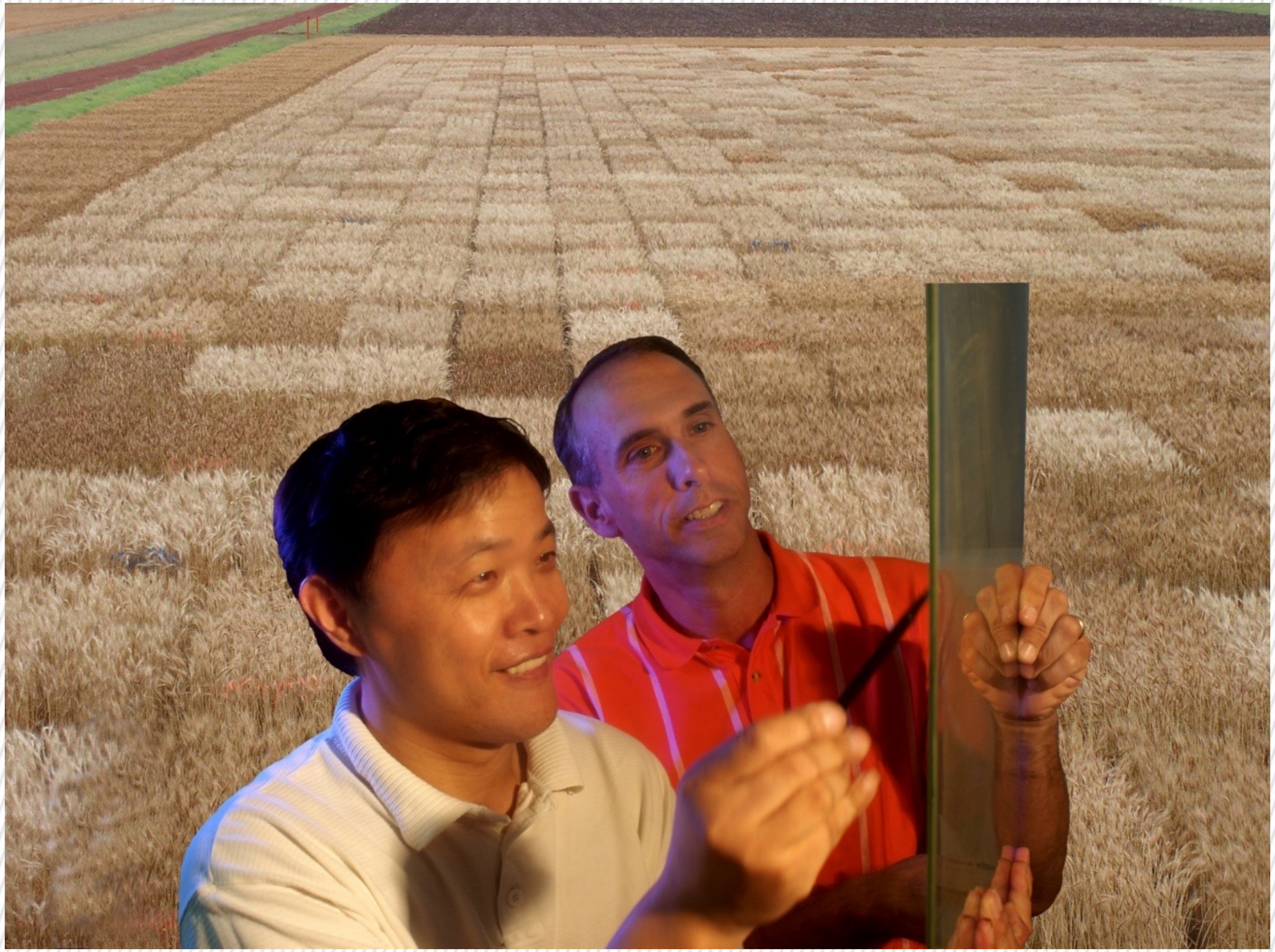
VRN2, **Yan et al. Science, 303:1640 -1644 (2004)**





VRN3, **Yan et al. PNAS, 103:19581-19586 (2006)**

Special acknowledgements



The cloned genes have been and will be utilized in new wheat varieties!!!

Winter wheat varieties released

- •2020 ‘Uncharted’
- •2020 ‘Strad CL Plus’
- •2020 ‘Breakthrough’
- •2020 ‘Butlers Gold’
- •2020 ‘Big Country’
- •2019 ‘OK Corral’
- •2018 ‘Showdown’
- •2018 ‘Green Hammer’
- •2018 ‘Bakers Ann’
- •2018 ‘Skydance’
- •2017 ‘Smith’s Gold’.
- •2017 ‘Spirit Rider’.
- •2017 ‘Lonerider’.
- •2015 ‘Bentley’.
- •2013 ‘Doublestop CL Plus’.
- •2012 ‘Iba’.
- •2012 ‘Gallagher’.
- •2011 ‘Ruby Lee’.
- •2011 ‘Garrison’.
- •2009 ‘Pete’.
- •2009 ‘Billings’.
- •2008 ‘Duster’.

Acknowledgments

TaCol-B5

Wheat Molecular Genetics Lab

Xiaoyu Zhang
Haiyan Jia
Tian Li
Jizhong Wu
Ragupathi Nagarajan
Lei Lei
Chia-Cheng Kan
Carol Powers

Wheat Improvement Team at OSU

Charles Chen
Brett F. Carver

Collaborators

Zhiyong Liu (CAS, China)
Wei Hua (CAS, China)

Q5A

Zhengqiang Ma (NJAU, China)

Lab members

Ragupathi Nagarajan
Wenxuan Zhai,
Jennifer Tapia,
Juan Carlos Yanez
Ryan Landy
Min Fan

Linyi Qiao
Yahya Rauf
Daijing Zhang
Shulin Xue
Genqiao Li
……

Dillon and Lois Hodges Professorship

Wheat CAP
Coordinated Agricultural Project

USDA NIFA

OCAST
Oklahoma Center for the Advancement of Science and Technology

Oklahoma Wheat Commission

liuling.yan@okstate.edu
LinkedIn @liuling-yan
@liuling_yan